Neural circuit mechanisms to transform cerebellar population dynamics for motor control in monkeys

David J. Herzfeld and Stephen G. Lisberger

Department of Neurobiology, Duke University School of Medicine, Durham, NC, 27710, USA

Correspondence: David J. Herzfeld, Department of Neurobiology, 311 Research Drive, Box 3209, Duke University School of Medicine, Durham, NC 27710, USA. Email: david.herzfeld@duke.edu

Running title: Cerebellar computations during smooth pursuit

Number of words: Abstract, 150; Introduction, 690; Discussion, 1,636; Total, 9,505

Keywords: flocculus, neuron type, smooth pursuit, learning, basis set, mossy fiber, unipolar brush cell, Golgi cell, granule cell, circuit computation

Number of figures: total, 7; color, 7

Abstract

We exploit identification of neuron types during extracellular recording to demonstrate how the cerebellar cortex's well-established architecture transforms inputs into outputs. During smooth pursuit eye movements, the floccular complex performs distinct input-output transformations of temporal dynamics and directional response properties. The responses of different interneuron types localize the circuit mechanisms of each transformation. Mossy fibers and unipolar brush cells emphasize eye position dynamics uniformly across the cardinal axes; Purkinje cells and molecular layer interneurons code eye velocity along directionally biased axes; Golgi cells show unmodulated firing. Differential directional response properties of different neuron types localize the directional input-output transformation to the last-order inputs to Purkinje cells. Differential temporal dynamics pinpoint the site of the temporal input-output transformation to granule cells. Specific granule cell population dynamics allow the temporal transformations required in the area we study and generalize to many temporal transformations, providing a complete framework to understand cerebellar circuit computation.

Impact statement

We dissect the circuit computations performed by the floccular complex of the cerebellum during an exemplar sensory-motor behavior, taking advantage of knowledge of the circuit architecture, existence of discrete neuron types, and a newfound ability to identify neuron types from extracellular recordings. Our results describe the contributions of the major neuron types to the cerebellar input-output computations, identify the population dynamics needed in granule cells to support those computations, and to create a basis set to enable temporally-specific motor behavior and motor learning.

Introduction

Understanding how neural circuits process information to generate behavior is a fundamental goal of systems neuroscience. Population dynamics[1,2] is a key feature of neural processing, and the challenge of generating behavior can be envisioned as multiple transformations[3] of population dynamics - from sensory inputs[4] to intermediate representations[5] to the ultimate drive of precise muscle activations[6]. A complete understanding of how neural circuits transform population dynamics requires four key components: 1) precise control over the input stimuli and quantifiable behavioral outputs, 2) fundamental knowledge of the necessary brain regions and their underlying neuron types and circuit architecture, 3) recordings of activity from complete neural populations in the requisite brain regions, and 4) the ability to link neuron types with neural responses for the major neuron types within each circuit. We have achieved all four components for one brain region and one sensory-motor behavior and therefore can explain how a specific circuit computes, in both space and time.

The cerebellum offers unique advantages for understanding neural circuit computations. First, the cytoarchitecture of the cerebellum is highly conserved across regions and species[7,8], with a well-characterized circuit organization[9] and discrete neuron types[9,10] organized around a primary three-layer feedforward network[11,12]: incoming mossy fibers synapse onto granule cells, whose parallel fiber axons provide excitatory input to Purkinje cells, the sole output of the cerebellar

cortex. Additional neuron types[10], including Golgi cells, unipolar brush cells, and molecular layer interneurons (among others[13,14]), use both feed-forward and recurrent connections[11] to modulate the activity within the pathway between mossy fibers and Purkinje cells. Second, many cerebellar regions are crucial for the precise execution of motor behaviors[15–19]. Motor behaviors provide an ideal behavior to study neural circuit function as the parameters of movements can be quantified with high spatial and temporal precision. Third, recent advances in large-scale multi-contact electrode technology[20,21] have enabled extracellular recordings from large populations of well-isolated cerebellar neurons during motor behaviors[22–27]. Finally, and crucially, our recent work allows identification of many of the cerebellar neuron types directly from extracellular recordings made during behavior[23,28]. Together, the cerebellar circuit, advances in recording technology, parametrizable motor behavior, and cerebellar neuron-type identification satisfy the four key requirements to link neural circuit function to behavior.

Here, we reveal how a full cerebellar circuit computes by taking advantage of knowing the input-output transformation in the cerebellar floccular complex during an exemplar motor behavior for which it is essential[18,29], smooth pursuit eye movements. The floccular complex causally controls smooth pursuit: it has disynaptic connections to extraocular motoneurons[30] and microstimulation elicits short-latency smooth eye movements[31,32]. Recordings from directly-identified Purkinje cells[33,34] and putative mossy fiber inputs tentatively characterize the input-output transformation of the floccular circuit. Purkinje cell responses are related to eye velocity[32,35,36] with directional preferences that align with the vestibular labyrinths[37]. Mossy fiber inputs exhibit responses resembling motor corollary discharge signals, driven principally by a combination of eye position and velocity[37,38] along the horizontal and vertical axes[37]. The requirement for seamless interaction of coordinate frames dictated by vestibular sensory signals and eye movement motor corollary discharge creates contrasting dynamics of floccular inputs and outputs and defines a set of floccular input-output transformations.

Our contribution is to show how specific interneuron types work together to transform mossy fiber inputs into Purkinje cell outputs through circuit function and to highlight how the floccular circuit operation could generalize widely. We overcame two challenges that have prevented a complete understanding of cerebellar computations from extracellular recordings to date: 1) granule cells are inaccessible to conventional recording technologies[23], and 2) Purkinje cells were the sole neuron type within the cerebellar circuit that could be identified definitively. Our recordings from several classes of identified interneurons allow us to discriminate among previously-proposed models of granule cell dynamics[39–44]. We used computational modeling to identify granule cell representations with the capacity to generate the measured responses of all downstream neural populations. The result is a complete model of cerebellar circuit processing that incorporates granule cell computations, accounts for downstream population dynamics, and also explains several previous behavioral and neurophysiological observations during cerebellar-dependent smooth pursuit learning. Thus, we believe we have identified how a complete neural circuit computes to generate well-calibrated motor behavior.

Results

We first confirm prior reports that the floccular complex performs directional and temporal transformations, rather than acting as a simple relay from mossy fiber inputs to Purkinje cell outputs. We further demonstrate that different functional computations are partitioned along neuron-type boundaries.

The floccular complex performs active computations during pursuit

We recorded extracellular action potentials from the floccular circuit (Figure 1A) while monkeys performed smooth pursuit eye movements. Using expert identification[28] (see ***Methods***), validated via ground-truth optogenetic identification in mice[23], we successfully identified all major neuron types schematized in Figure 1A, with the exception of granule cells. During smooth pursuit tracking, monkeys fixated a central dot for a randomized interval (400-800 ms) before the target started to move at a constant velocity for 650 ms. In the example trial in Figure 1B, the target moved rightward at 20 deg/s. Monkeys tracked the target throughout its motion and maintained eccentric fixation for an additional 350 ms to receive fluid reward. Targets moved at 10, 20, or 30 deg/s along the four cardinal directions.

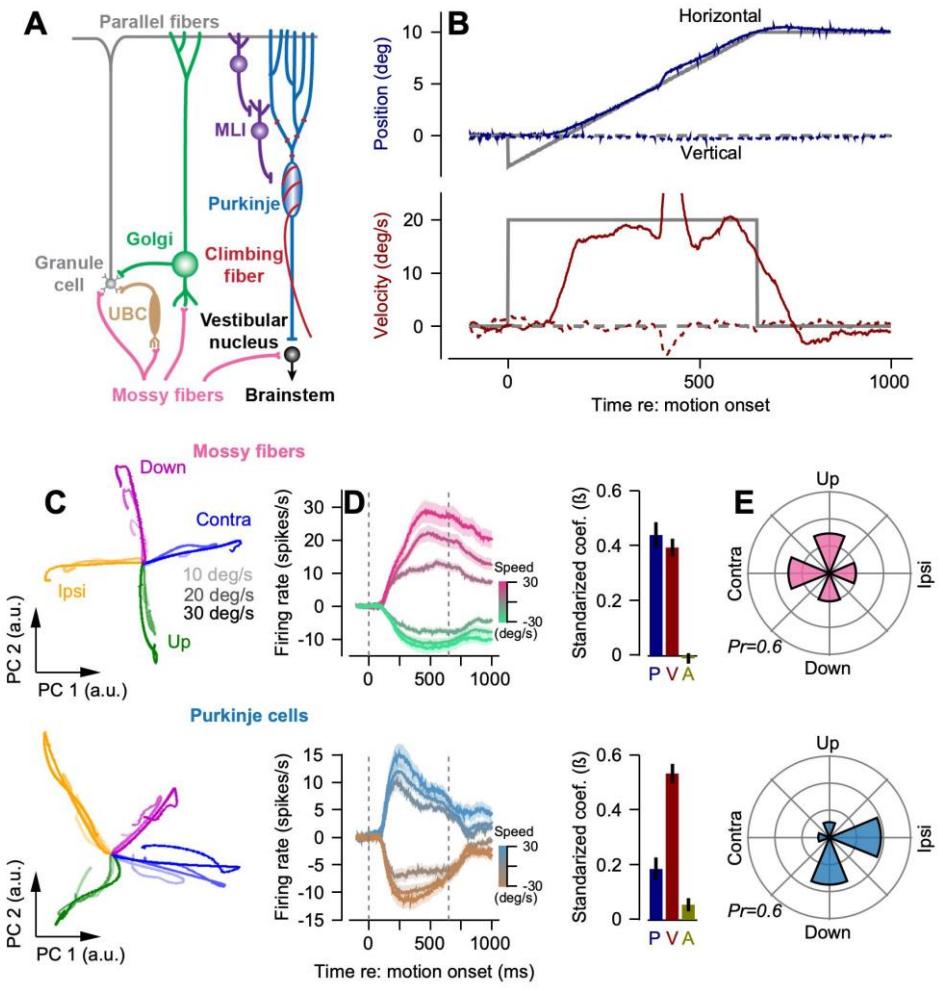

Figure 1. Temporal and directional input-output transformations performed by the cerebellar floccular complex during smooth pursuit. (A) Simplified schematic of the cerebellar circuit in the floccular complex. (B) Eye and target position (top) and velocity (bottom) versus time for an example smooth pursuit trial where the target moved exclusively in the horizontal direction. (C) Population dynamics derived via principal component analysis conducted across target speeds (10, 20, 30 deg/s, increasingly dark lines) and directions (ipsiversive, contraversive, up, and down) for mossy fibers (top) and Purkinje cells (bottom). (D) Mean population responses of mossy fibers and Purkinje cells in their

preferred and anti-preferred directions, relative to their respective baseline responses. Shaded bands represent mean ± SEM across neurons. Histograms on the right summarize the regression coefficients for model of firing rate in the preferred and anti-preferred directions as a function of eye position (P), velocity (V), and acceleration (A). (E) Polar plots showing the distribution of preferred directions across the populations of recorded mossy fibers and Purkinje cells. Preferred directions are defined as pursuit direction with maximal mean positive deviation of firing rate (0 to 650 ms after target motion onset) relative to the pre-trial baseline response. We note that the difference in direction tuning in Purkinje cell state-space trajectories between C and E is due to the sign-agnostic nature of principal component analysis along with the comparable magnitudes of Purkinje cell modulation in the preferred and anti-preferred directions.

Separate principal component analyses of the firing rates in populations of mossy fibers and Purkinje cells reveal distinct neural state-space trajectories[1,45] (Figure 1C). Mossy fiber trajectories aligned closely with the horizontal and vertical axes of the eye. The trajectories terminated eccentrically in state-space with magnitudes that scaled with pursuit speed, indicating persistent activity related to eccentric eye positions across all speeds and directions following pursuit termination. In contrast, Purkinje cell trajectories showed principal axes that were rotated relative to the cardinal axes. Their neural trajectories returned to near the origin at movement completion, consistent with the absence of a relationship to eye position. For both neuron types, the first two dimensions captured a substantial portion of the response variance (38% and 29% for mossy fibers and 23% and 20% for Purkinje cells, a considerable amount given that we are accounting for responses in 4 directions by 3 speeds of target motion).

To directly link neural trajectories to firing patterns, we compared mean population responses in the preferred and anti-preferred directions (preferred + 180°) across all speeds (Figure 1D). Mossy fiber responses resembled eye position traces, with persistent firing at movement termination that scaled approximately linearly with pursuit speed (preferred direction $R^2 = 0.83 \pm 0.03$). The asymmetry of mossy fiber responses in the preferred versus anti-preferred directions stems from complete cessation of firing in the anti-preferred direction for some mossy fibers. In contrast, Purkinje cell responses peaked approximately 200 ms after motion onset with magnitudes that scaled with pursuit speed (preferred direction $R^2 = 0.75 \pm 0.03$) and returned to baseline at the end of pursuit regardless of final eccentric eye position. Thus, an important signature of whether neurons have position-related responses is whether firing rate (1) remains elevated at the end of the movement while the monkey fixates an eccentric target and (2) post-movement responses scale with target speed because faster target motions make the final eye position more eccentric. Regression analysis confirmed equal relations to eye position and velocity with no relation from eye acceleration for mossy fibers, and a dominant relation to eye velocity in Purkinje cells (histograms on right of Figure 1D). Statistical analysis verified that encoding of position ($t(185)=4.08$, $p < 10^{-5}$), velocity ($t(185)=-2.86$, $p=0.005$) and acceleration ($t(185)=-2.12$, $p=0.035$) all differed across mossy fibers versus Purkinje cells (independent samples t-test). Documentation of the responses of each individual Purkinje cell and mossy fiber appears in Supplemental Figure 1.

Directional tuning properties also differed between mossy fibers and Purkinje cells (Figure 1E). Mossy fibers showed uniformly distributed preferred directions across the principal axes (circular dispersion: $\overline{R} = 0.13$). Purkinje cells' preferred directions were more strongly biased,

with preferred directions close to ipsiversive and downwards ($\overline{R}$= 0.58), consistent with prior reports[22,35,36,38,46,47]. Together, these analyses reveal two fundamental circuit computations: a *temporal* transformation that converts the sustained activity of mossy fibers into transient responses in Purkinje cells and a *directional* transformation that converts uniformly distributed directional inputs into biased directional outputs.

Distinct temporal dynamics of different cerebellar neuron types

Having established that the cerebellum must perform transformations of mossy fiber population dynamics, our next goal was to identify whether different cerebellar neuron types have distinct computational roles in transforming inputs to outputs. Indeed, mean firing rate profiles for each neuron type during preferred direction pursuit at 20 deg/s suggested functional distinctions across neuron classes (Figure 2A). The responses of mossy fibers and unipolar brush cells closely tracked eye position and remained elevated during eccentric fixation at the end of pursuit. Molecular layer interneurons and Purkinje cells exhibited more transient temporal patterns related to eye velocity, with firing rates that returned to close to baseline at the end of the movement while the monkey fixates an eccentric target. Golgi cells showed minimal modulation during pursuit.

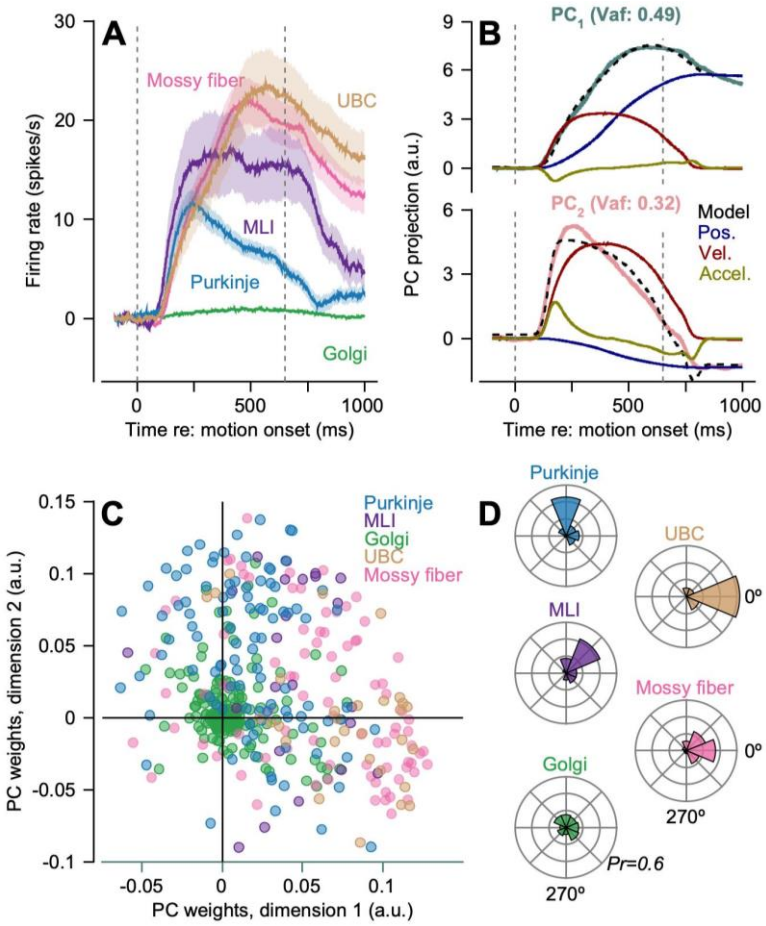

Figure 2. Functional heterogeneity of responses across cerebellar neuron types during smooth pursuit eye movements. (A) Mean modulation of firing rate as a function of time for cerebellar

neurons, segregated by neuron type, during 20 deg/s pursuit in each neuron's preferred direction. Firing rates show the modulation relative to a pretrial baseline. Error bands show mean ± SEM across neurons. (B) Top and bottom plots show the projection of neuron firing rates in their preferred direction onto the first two principal components (PC_1 and PC_2) computed across all identified neurons recorded in the cerebellar flocculus. The different colored traces show the time course of the relevant principal component projection, and the contributions of eye position, velocity, and acceleration derived by fitting a kinematic model to the projected timeseries. Black dashed trace shows the overall fit of the kinematic model to the time course of the projected principal component. (C) Scatter plot of the weighting onto the first two principal components. Each dot corresponds to an individual identified neuron, colored by neuron type. Mean centroids are: Golgi cells, (0.006 ± 0.002, 0.005 ± 0.002); Purkinje cells, (0.016 ± 0.003, 0.042 ± 0.005); molecular layer interneurons (0.042 ± 0.007, 0.034 ± 0.013); mossy fibers (0.059 ± 0.005, 0.014 ± 0.006); and unipolar brush cells (0.075 ± 0.007, 0.001 ± 0.008). (D) Distribution of angular locations, irrespective of magnitude, for each neuron type derived from the scatter plot shown in C.

Functional segregation across neuron types remained clear when analyzed at the level of individual neurons instead of population means. Principal component analysis of trial-averaged responses for pursuit in each neuron's preferred direction at 20 deg/s identified two dominant components that together explained 81% of the population variance (49% and 32%, respectively). The first component (Figure 2B, top) showed a near monotonic increase over time with sustained deviation from baseline following pursuit termination. Linear regression of kinematic variables showed a strong relation to eye position (standardized coefficient, $\beta = 0.84$) and eye velocity ($\beta = 0.50$). The second principal component (Figure 2B, bottom) exhibited an early peak followed by a return to near baseline; it was strongly related to eye velocity ($\beta = 0.83$) and somewhat to eye acceleration ($\beta = 0.24$).

Individual neuron weights in the two-dimensional space defined by the two dominant components revealed across-class distinctions along with some within-class variation (Figure 2C). For instance, Golgi cells are concentrated near the origin, Purkinje cells show near zero weights in the first dimension but consistently positive weights in the second dimension, and mossy fibers and unipolar brush cells show the opposite trend with weights centered near zero in the second dimension but strongly positive in the first dimension. Permutation analysis revealed that all neuron type centroids were significantly different from a neuron-type agnostic distribution (mossy fibers, $n = 86$, $p = 10^{-4}$; unipolar brush cells, $n = 30$, $p = 10^{-4}$ Golgi cells, $n = 186$, $p = 10^{-4}$; Purkinje cells, $n = 101$, $p = 10^{-4}$; molecular layer interneurons, $n = 23$, $p = 0.039$). Much of the within-neuron-type variation in the spatial distribution shown in Figure 2C results from heterogeneity in response magnitude rather than fundamental differences in temporal processing within neuron types. Probability distributions of the angle of each neuron within the two-dimensional space (Figure 2D) show marked differences across neuron types in the distributions of the relative contributions of the first two principal components to neuron firing. The distinctive neuron-type-specific response patterns highlight active computation within the circuit to alter the temporal properties of the signals, with potentially distinct computational roles for each neuron type.

Distinct roles of cerebellar interneurons in floccular circuit processing

Using our ability to identify other neuron types in the cerebellar circuit reliably[23,28], we next examine the contributions of Golgi cells, molecular layer interneurons, and unipolar brush cells

192 to the input-output transformations as these neuron types likely serve to alter information
193 transmission through the cerebellar circuit.

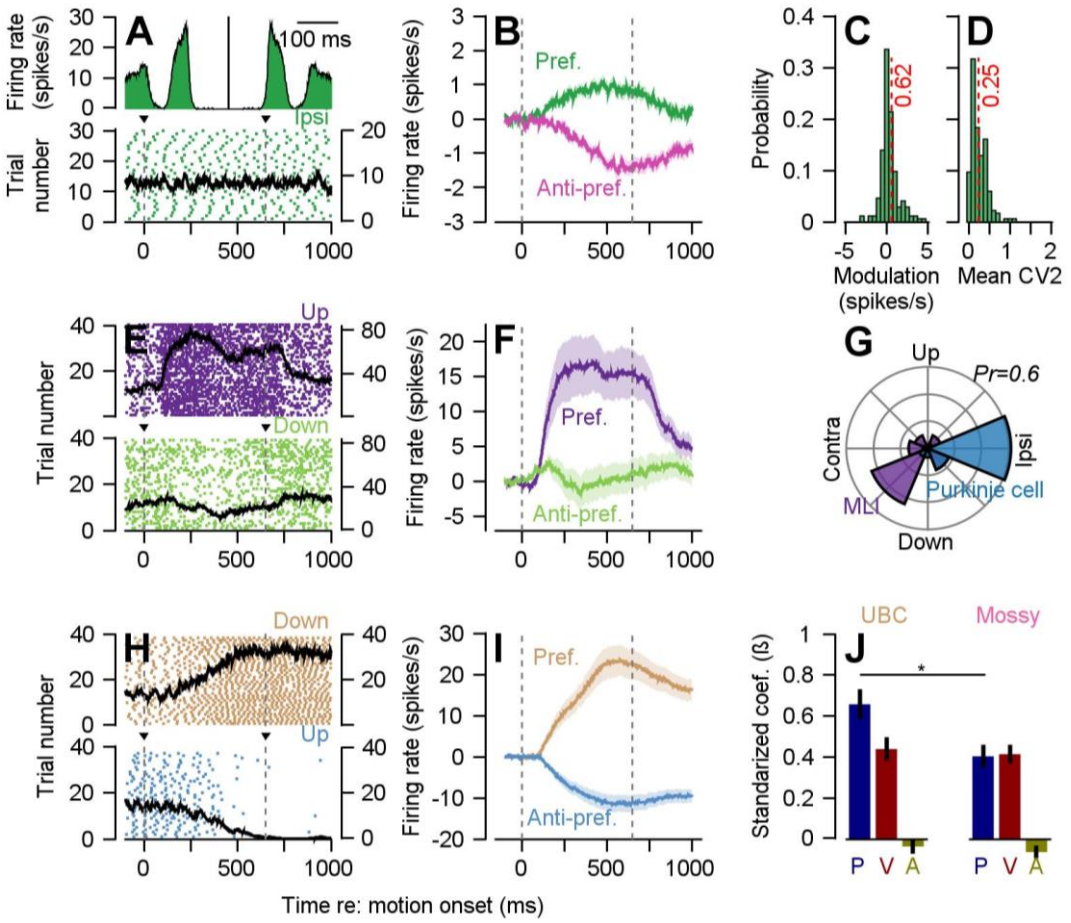

Figure 3. Response characteristics of Golgi cells, molecular layer interneurons, and unipolar brush cells during smooth pursuit eye movements. (A) Autocorrelogram of an exemplar Golgi cell (top). Raster plot of the same Golgi cell aligned to the onset of ipsiversive target motion (bottom). Black curve shows the mean firing rate across all trials shown in the raster plot. (B) Preferred and anti-preferred (preferred + 180°) direction modulation of firing, averaged across the complete Golgi cell population, relative to baseline firing. (C) Probability distribution showing the modulation across the Golgi cell population in the preferred pursuit direction. (D) Probability distribution of CV2 of Golgi cells computed across complete recording sessions. Red dotted lines in C and D denote population means across all Golgi cells. (E) Raster plots for an exemplar molecular layer interneuron in its preferred (top, upwards pursuit) and anti-preferred directions (bottom, downwards pursuit). Black curves denote the mean firing rate of the molecular layer interneuron across the trials shown in the raster. (F) Mean responses of the complete molecular layer interneuron population in the preferred and anti-preferred pursuit directions (20 deg/s pursuit), again relative to baseline firing. (G) Probability distributions of preferred directions for molecular layer interneurons (purple) and simultaneously-recorded Purkinje cells (blue). (H) Rasters for an exemplar unipolar brush cell in its preferred (downwards) and anti-preferred (upwards) pursuit directions. Black curves denote mean firing rates during pursuit across the trials plotted in the raster. (I) Mean firing rate of the unipolar brush cell population, relative to baseline, in their preferred and anti-preferred directions. (J) Comparison of

standardized coefficients relating eye position (P), velocity (V) and acceleration (A) to firing rates of mossy fibers (pink, top) and unipolar brush cells (brown, bottom) for pursuit in each neuron's preferred direction at 20 deg/s. Shaded regions and error bars in all panels denote mean ± SEM across neurons.

Golgi cells are inhibitory neurons that are hypothesized to regulate granule cell activity via recurrent excitation from granule cells' parallel fibers[48–50]. Given their position in the cerebellar circuit, Golgi cells could contribute to either the temporal or directional transformations. For instance, Golgi cell responses consistent with eye position might subtract or block the component of mossy fiber activity related to eye position via feedforward or recurrent inhibition, thereby creating granule cell responses related primarily to eye velocity. Yet, we found that Golgi cells in the flocculus complex showed minimal modulation during smooth pursuit.

A typical Golgi cell showed regular firing at a low rate of ~12 spikes/s and no temporal modulation of firing during ipsiversive pursuit at 20 deg/s (Figure 3A). The full Golgi cell population also showed limited modulation during pursuit in their preferred or anti-preferred directions at 20 deg/s (Figure 3B), quantified as the mean modulation in the preferred direction (Figure 3C, 0.62 ± 0.13 spikes/s). Golgi cell firing was highly regular as measured by the mean "CV2[51]" (Figure 3D, 0.25 ± 0.01). Only 46/186 Golgi cells showed statistically significant modulation via permutation testing across pursuit directions. We conclude that Golgi cells provide tonic but temporally-unmodulated inhibition of granule cells.

Molecular layer interneurons are positioned to perform feedforward or lateral inhibition[52] of Purkinje cells and thus could impact Purkinje cell activity both directionally and temporally. We identified molecular layer interneurons by their monosynaptic inhibition of known Purkinje cells in spike-timing cross-correlograms (see ***Methods***). In its preferred direction, a typical molecular layer interneuron's firing (Figure 3E) is related primarily to eye velocity, as it returned to near baseline levels at pursuit termination, mirroring Purkinje cells. The same molecular layer interneuron displayed minimal modulation of firing rate during pursuit in the anti-preferred direction, a trend confirmed in averages of firing rate across the complete molecular layer interneuron population (Figure 3F). The mean modulation of the molecular layer interneuron population was 12.2 ± 3.0 spikes/s in their preferred direction, but only 0.4 ± 1.6 spikes/s in their anti-preferred direction. The preferred direction of molecular layer interneurons was almost always opposite that of their connected Purkinje cells (Figure 3G, mean angular distance: -121.8° ± 10.4°, circular mean ± SEM). The opposite preference for pursuit across the two populations suggests that molecular layer interneurons likely contribute to the directional preferences of Purkinje cells by providing inhibition in the Purkinje cell's anti-preferred direction of pursuit.

Unipolar brush cells are relatively common in the oculomotor regions of the cerebellar cortex, but relatively rare in other cerebellar regions[53], suggesting a potential specialization for oculomotor behavior[54]. A typical unipolar brush cell (Figure 3H) showed clear positive and negative modulation in its preferred and anti-preferred directions and had fully sustained modulation in both directions after pursuit termination implying that they, like mossy fibers, respond primarily to eye position. The same response profile appeared in the averages across the complete population of unipolar brush cells, with sustained activity post-pursuit (Figure 3I). Regression analysis to fit the firing rate of each neuron as a function of the simultaneously recorded eye position, velocity, and acceleration showed stronger eye position encoding for

unipolar brush cells compared to mossy fibers (Figure 3J, standardized coefficient, $\beta = 0.65 \pm 0.07$ versus 0.40 ± 0.06, $t(114) = -2.36$, $p = 0.02$). The more dominant eye position response of the full population of unipolar brush cells suggests that they either perform integration of their mossy fiber inputs, consistent with previous *in vitro* analyses[55,56], or selectively discount velocity signals in their inputs from mossy fibers. Documentation of the response of each individual Golgi cell, molecular layer interneuron, and unipolar brush cell appears in Supplemental Figure 1.

Golgi cells, molecular layer interneurons, and unipolar brush cells do not seem to play a role in the temporal input-output transformation in the floccular complex. Golgi cells are largely unmodulated, molecular layer interneurons are already related to eye velocity, and unipolar brush cells are more related to position than are mossy fibers. Any role for the excitatory unipolar brush cells would require subtraction from spontaneous activity that is thought not to exist in granule cells[57]. Finally, the temporal transformation cannot be explained by synchrony in multiple mossy fibers inputs to a granule cell because we found no evidence for mossy fiber synchrony other than the amount expected from covariation of mossy fiber firing rates during pursuit (Supplemental Figure 2).

A data-driven approach for identifying granule cell transformations

Two key findings motivated a computational investigation of the transformation of mossy fiber input in the granule cell layer, given that we were not able to record from granule cells themselves. First, our analysis of other cerebellar neuron types, above, failed to reveal any candidates to mediate the temporal transformation from mossy fiber position signals to the velocity-related Purkinje cell responses. Second, Purkinje cell and molecular interneuron population responses both reflected the absence of a large position component, suggesting that they receive common inputs after a temporal transformation in the granule cell layer.

We started with a modeling approach using recurrent neural networks (Figure 4A), with the goal of identifying the requisite properties of granule cell processing in a data-driven manner. Recurrent neural networks are particularly amenable to understanding circuit computations because the dynamics governing how the model achieves the desired computation is unconstrained yet amenable to rigorous interrogation[58–60]. We realize that the recurrent neural network is not biologically realistic, but our goal in this analysis was to ask a computational question about 'how' to solve the computational problem and defer realistic biological implementation until we have identified the crucial components of the transformation.

We fed single-trial mossy fiber inputs into a recurrent "hidden" layer whose weighted outputs were then supplied as inputs to model "granule units" (Figure 4A, arrowhead). The granule units were constrained to have strictly positive responses by a sigmoid activation function. Their outputs were transformed via a fully-connected non-negative weight matrix that was optimized to predict the time-varying single-trial firing rates of Purkinje cells (see ***Methods***). The recurrent neural network in the model produced dynamics that allowed the model Purkinje cells to reproduce the mean temporal profiles of recorded Purkinje cell responses across a range of pursuit speeds (Figure 4B, $R^2 = 0.97$).

We obtained insights into the underlying computations by examining the diversity of temporal response properties exhibited by the granule units. When the recurrent neural network was

supplied with a step input, some units produced transient, high-pass filter like responses, while others showed more sustained, low-pass filter like activity (Figure 4C), reminiscent of bandpass filters with varying time constants. The latency of each granule unit's peak response was correlated to its response duration (correlation test between rise and decay time constants, $t(70) = 6.88$, $p < 10^{-9}$). Similar patterns of units that we could characterize as high-pass and low-pass were evident when we drove the model with trial-averaged mossy fiber pursuit responses (Figure 4D). Thus, the model network learned a range of temporal filtering characteristics where the degree of filtering and subsequent duration of firing were temporally linked.

A generative model of granule cell temporal transformations

We developed a generative model of granule cell firing that could capture the key characteristics revealed by the recurrent neural network model in Figure 4A-D. Replacing the recurrent neural network model with a generative model served two primary purposes. First, it allowed us to embed granule cell population responses in otherwise biomimetic circuit models. For example, we could simulate the fact that granule cells receive only four upstream inputs[12,61,62], a feature not present in the recurrent network model. Second, a generative model affords the flexibility to predict granule cell population dynamics to novel inputs, enabling us and others to extend the model and investigate how it generalizes beyond the specific behavior and dataset from our cerebellar region and task.

Based on the predictions of the RNN model and from iterative attempts to understand how the granule cell population response would contribute to multiple features of pursuit behavior, we incorporated four key features into the generative model: 1) a broad range of temporal filter responses across the population so that granule units temporally tile the duration of the trial; 2) strongly correlated onset and offset time constants to enable scalar variability[63] (i.e., "Weber law") behavior for learned timing; 3) increasing granule unit population activity with increasing pursuit speeds to allow Purkinje cell firing to scale with speed; and 4) consistent timing of individual granule unit activations across pursuit speed to enable generalization across pursuit speeds. These characteristics are consistent with previously proposed "spectral timing" models of granule cell activity[39,64,65].

We simulated each granule unit as a dynamical system with two states. The system functioned as a generalized differentiator whose output was half-wave rectified (see ***Methods***). The resulting granule units exhibited a range of temporal filtering properties that agreed well with the step-responses of model granule units predicted by the recurrent neural network. Some had rapid input integration and fast decay leading to short-latency, transient responses, while others had slower integration and decay leading to more prolonged activity with later peaks (Figure 4E). We contrived a strong correlation between the rise and decay time constants ($t(88) = 31.1$, $p = 10^{-49}$).

The granule unit responses in the generative model showed temporal diversity similar to that in the recurrent neural network model (compare Figures 4D and F) when supplied with the mean mossy fiber response during pursuit (Figure 4G) as input. As expected, the activity of early-peaking units returned to baseline quickly and that of late-peaking units remained elevated for longer durations (see responses of three exemplar granule units, Figure 4H). The mean response of a population of 100 granule units (Figure 4I) qualitatively matched the temporal profiles observed in experimentally recorded Purkinje cells (Figure 4J, blue curves). Further, optimizing the weights of the granule units via non-negative least squares to the mean responses across the

population of Purkinje cells provided an excellent fit (Figure 4J, R^2=0.92). Thus, temporal tiling in a population of granule cells can solve the position-to-velocity transformation from mossy fiber to Purkinje cells. We note that there are a number of implementations of short-term plasticity in the mossy fiber to granule cell synapse that could produce the temporally-tiled granule cell basis set implemented by our models of granule unit activity[66]. Also, other extant models of granule cell responses[42] share some of the features of the generative model and, therefore, are able to reproduce the time varying Purkinje cell firing to some degree, albeit less well than our model (Supplemental Figure 3).

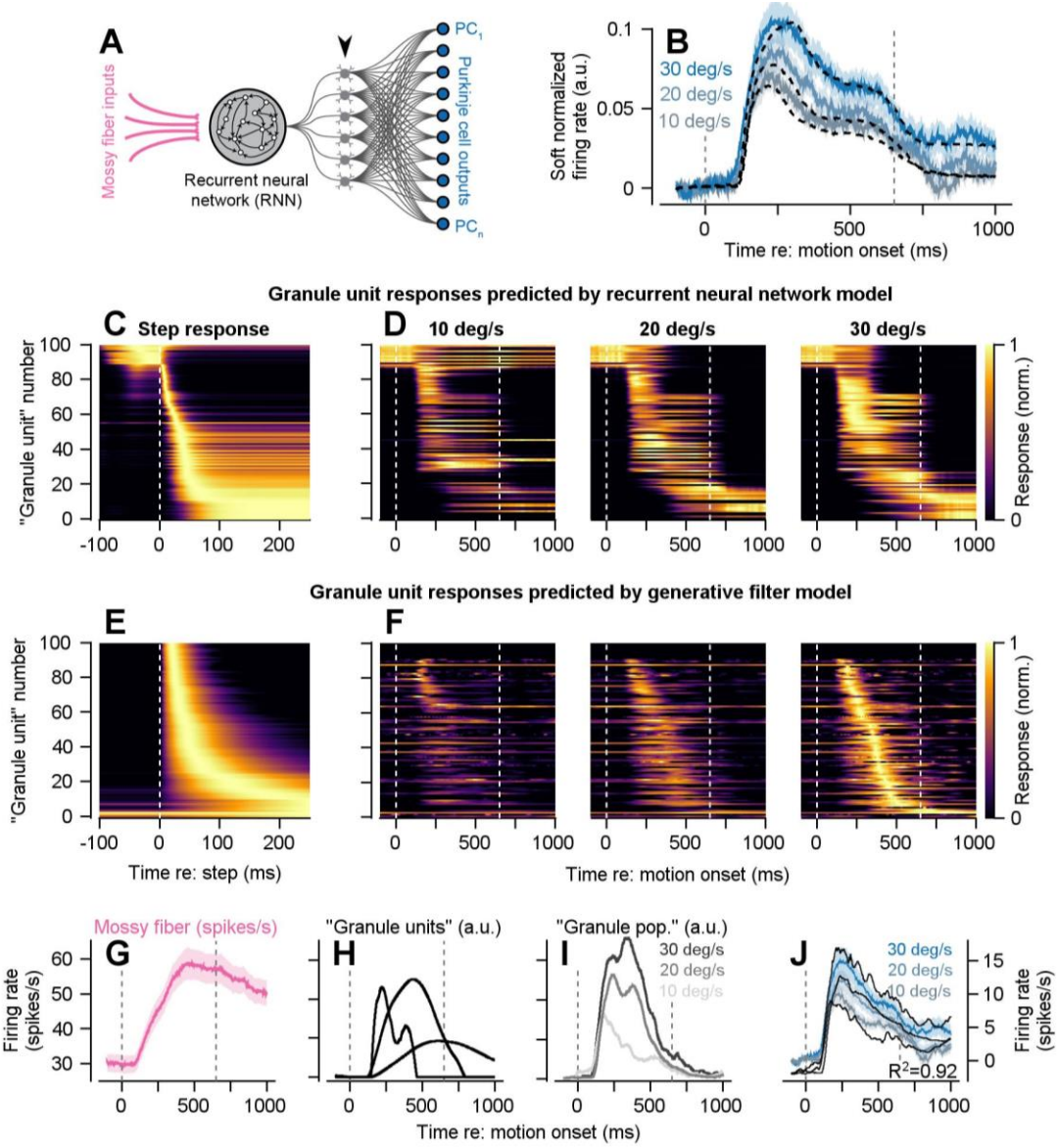

Figure 4. Models reveal a potential mechanism of the temporal transformation between mossy fibers and Purkinje cells. (A) Recurrent neural network (RNN) model architecture schematic. RNN units acted as a pool of potential dynamics for putative "granule units" (black arrowhead). (B) Success of the model architecture shown in A to account for Purkinje cell data. Dashed black traces show mean

RNN model responses to three different speeds of target motion. Blue traces show soft-normalized[45] population mean Purkinje cell firing rates across speeds. (C) Responses of granule units in the RNN model to a step change in the input from mossy fibers at *t=0* ms. Each unit's response was normalized to its maximum firing rate. Rows are ordered by the time of the peak response. (D) Responses of "granule units" in the RNN model to the three trained pursuit speeds shown in B. Each granule unit's response was normalized to its peak response across all three pursuit speeds and ordered based on the timing of the peak response to 30 deg/s target motion (far right). (E) Responses of granule units in the generative model to a step input of mossy fiber responses, as in C. (F) Responses of granule units in the generative model to four randomly sampled mossy fiber inputs across pursuit speeds. Granule unit responses are normalized and ordered based on the time of their peak response to 30 deg/s pursuit, as in D. (G) Mean responses of mossy fibers supplied as input to the generative granule unit model. (H) Representative responses of 3 granule units to the mean mossy fiber input shown in G. (I) Mean population responses of the granule unit populations in F across pursuit speeds. (J) Fits (black traces) to the mean responses of Purkinje cells across pursuit speeds (colored traces) by the non-negative weighted outputs of the generative model of granule units.

A cerebellar circuit model that reproduces the directional and temporal response properties of individual Purkinje cells and molecular layer interneurons

Armed with a generative model of granule population processing, our next step was to determine whether weighted combinations of generative granule cells could reproduce the temporal firing rate responses and direction tuning of individual molecular layer interneurons and Purkinje cells during pursuit eye movements. The circuit model (Figure 5A) consisted of 86 mossy fibers with the firing profiles from our recorded population, 1,000 granule units based on the generative filter model, 250 molecular layer interneurons (augmented from our recorded *n=23* population), and our 101 recorded Purkinje cells. Each granule unit received input from 4 randomly-chosen mossy fibers (Figure 4B, left top), consistent with the number of inputs observed anatomically[12,61,62]. As the mossy fiber population showed preferred direction distributions aligned to the cardinal axes with approximately equal magnitudes (Figure 5B, right top) and relatively symmetric tuning (Figure 5B, left bottom), the modeled granule unit population inherited the uniform distribution of preferred directions across the cardinal directions (Figure 5B, right bottom). We used non-negative least squares regression of the full population of granule unit inputs to fit the firing profiles of the 250 molecular layer interneurons across pursuit directions (see ***Methods***). We obtained excellent fits of both direction tuning and temporal dynamics (top row of Figure 5C, purple trace, R^2=0.98) with a mean R^2 of 0.98 across the augmented population of molecular layer interneurons (Figure 5E).

We also obtained excellent fits to the simple-spike firing of all individual Purkinje cells (examples in Figure 5C) with weighted inputs from both the granule cell activations and the augmented molecular layer interneuron population; preferred directions of molecular layer interneurons were rotated by 180° to account for the opposite preferred directions relative to Purkinje cells (Figure 3G). For the three example Purkinje cell responses in Figure 5C the minimum R^2 was 0.97. Across the population of 101 Purkinje cells, the mean R^2 was 0.94 (Figure 5F). Thus, the simplified model of granule cell dynamics provides a sufficient basis set to account fully for the activity of individual molecular layer interneurons and Purkinje cells.

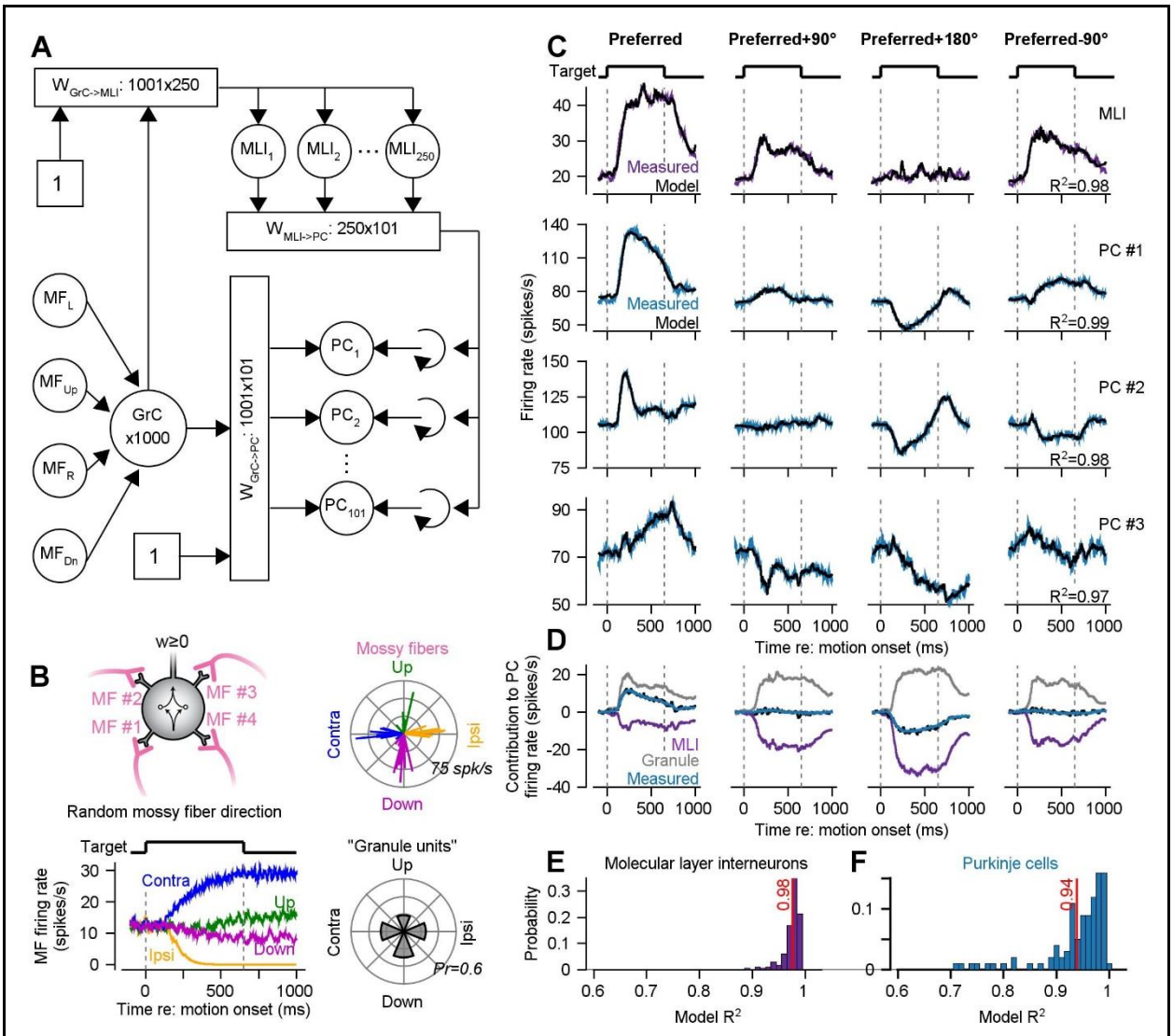

Figure 5. Model granule cell populations provide a sufficient basis for simulation of molecular layer interneuron and Purkinje cell responses. (A) Schematic diagram of a circuit model of the cerebellar cortex used to predict the responses of individual molecular layer interneurons and Purkinje cells. (B) Schematic denotes random convergence of four mossy fibers on a model granule cell (with intrinsic dynamics, see ***Methods***). Bottom plot shows the responses of an example mossy fiber to 20 deg/s pursuit in the four cardinal directions. Right polar plots show the magnitude and preferred directions of mossy fibers (top) and directional preferences of modeled granule cell units (bottom). (C) Fitted responses of an exemplar molecular layer interneuron (top row) and three exemplar Purkinje cells (rows 2-4) to 20 deg/s pursuit in four directions relative to each neuron's preferred direction of pursuit. Black traces show model fits in all panels. (D) Blue traces show mean modulation of Purkinje cell firing, black graces show mean model fits, and purple and gray graces show mean inputs from molecular layer interneurons and granule units. (E) Mean model R^2s for molecular layer interneurons and (F) Purkinje cells. Vertical red lines in E-F denote the mean across the respective populations.

To understand how the model works, and potentially how the cerebellar circuit computes, we analyzed the weighted contribution of granule cells and molecular layer interneurons to Purkinje cell firing across directions (Figure 5D). In the preferred direction, Purkinje cell activity was

driven almost entirely by excitatory input from the granule units. In the anti-preferred direction, the molecular layer interneuron population provided significant inhibitory input to counteract the fact that granule units delivered increased rather than decreased excitation. In orthogonal axes, inputs from granule units and molecular layer interneurons approximately cancelled, resulting in Purkinje cell responses that were unmodulated from baseline levels. Thus, our model predicts that the two inputs operate partly as reciprocal excitation and inhibition and partly in the balanced E/I regime that seems to predominate in the cerebral cortex[67–70].

Emergent properties of the cerebellar circuit model

Our cerebellar circuit model was capable of performing the temporal and directional transformation between mossy fiber inputs and Purkinje cell simple-spike firing during pursuit eye movements across directions. We next sought to test whether the model had emergent properties that could account for previously described behavioral and neurophysiological observations obtained during pursuit-direction learning[46,71–74].

Previous recordings in the floccular complex during pursuit learning tasks have demonstrated the critical role of climbing fiber-mediated plasticity at the parallel fiber to Purkinje cell synapse[46,47,73,75]. Complex-spike firing on single learning trials causes a well-time depression of simple-spike firing on the next trial[46,47,73,75]. Further, the probability of complex-spike firing across blocks of tens to hundreds of learning trials is correlated with the magnitude of reductions in simple-spike responses[46,75]. When imbued with these features, the circuit model in Figure 5 reproduced multiple experimental observations that were not built into its design. We will show below that the properties of the granule unit basis set in relation to time and eye velocity are necessary features for the emergent properties of the model during pursuit learning.

To produce directional learning in pursuit, the monkey tracks a target that moves at a constant speed in an initial "pursuit direction." After a fixed duration, an orthogonal velocity component (the "instruction") is added to the target's motion, so that the target moves diagonally[71,72] (Figure 6A, top). In the velocity domain, target motion comprises a 650-ms duration step of velocity in the initial, pursuit direction (Figure 6A, middle), along with a 400-ms duration step of velocity in the learning direction starting 250 ms later (Figure 6A, bottom). The instructive change in target direction drives behavioral learning, assayed in the next trial or over short bouts of learning trials. Learning comprises an appropriately-timed deviation of the eyes in the direction of the preceding instruction that accumulates over trials (Figure 6A, arrowhead). The instruction also evokes complex-spike responses in the 100 milliseconds following the instructive stimulus (Figure 6D, shaded region) and causes a well-timed depression of simple-spike firing rate (Figure 6G).

The model reproduces our previous results on generalization of single-trial learning[72]. To study generalization, we induced learning with pursuit target motion at 20 deg/s and measured learning in the subsequent test trial with slower or faster pursuit target motion (Figure 6B). In our experiments[72], eye movement deviation in the instruction direction of the test trial scaled with test target speed in the pursuit direction (Figure 6E). We simulated the learning conditions in our model by reducing the connection weights from the subset of parallel fibers active at the time of instruction during pursuit at 20 deg/s (250 ms, red arrow) and measuring the learning-induced changes in Purkinje cell simple-spike output on the subsequent trial for pursuit speeds of 10, 20, and 30 deg/s. The model predicted a learned depression in the Purkinje cell population response

relative to simulated probe trials without a preceding learning trial that scaled with pursuit target speeds (Figure 6H), in agreement with our behavioral results. The scaling of firing rates with pursuit speed arises because the magnitude of granule cell population activity increases with pursuit speeds and a fixed amount of synaptic depression causes a post-synaptic response that scales with the firing of the pre-synaptic fibers.

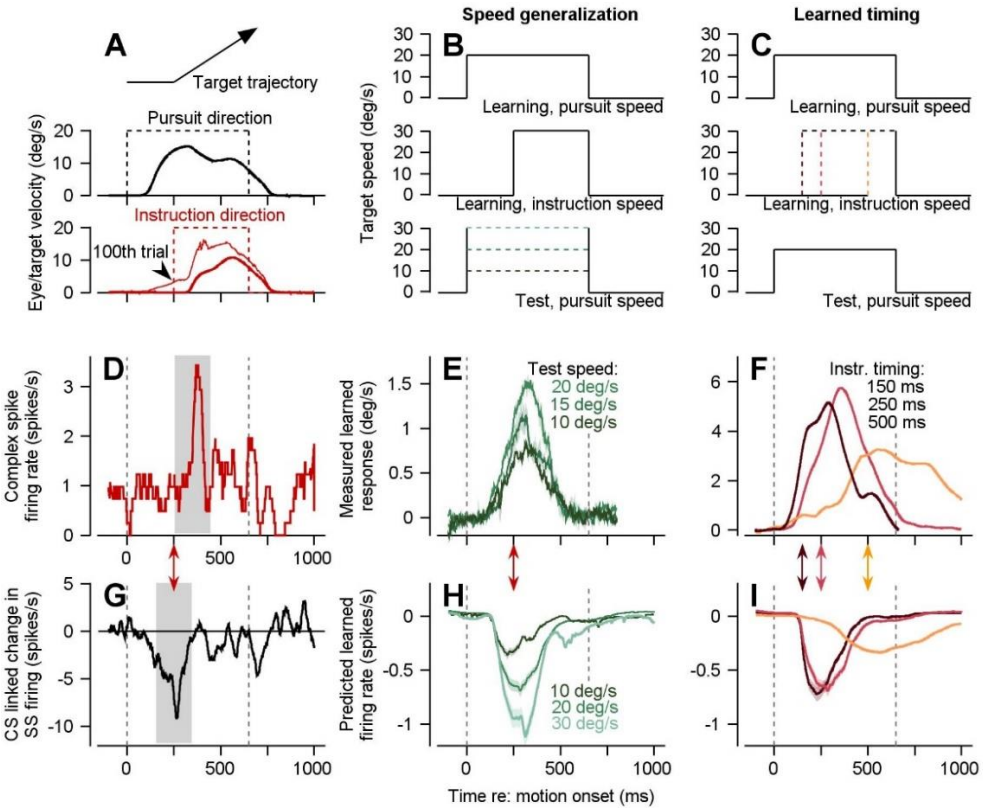

Figure 6. Emergent properties of the circuit model of the cerebellar cortex during pursuit learning. (A) Target position trajectory during a direction-learning trial (top). Eye and target velocity timeseries in the original pursuit direction (middle) and orthogonal instructive direction (bottom). Arrowhead in the bottom plot highlights eye velocity after 100 learning trials in the same direction. (B) Target velocity profiles for a learning paradigm that tests speed generalization. Learning is instantiated as in A (top and middle traces). Speed generalization is probed in test trials (bottom) where target speed in the original pursuit direction is varied. (C) Target velocity profiles for a pursuit learning paradigm that tests learned timing of the instructive stimulus. An instructive signal occurs at either 150 ms, 250 ms, or 500 ms after target motion onset (middle). Learning is probed after 100 learning trials in a test trial with target motion only in the pursuit direction (bottom). (D) Complex spike response measured during direction learning in trials where the instruction was in the preferred complex-spike direction (CS-on) of the Purkinje cell under study. Grey shaded region denotes instruction-linked complex-spike period. (E) Eye velocity responses after learning, tested with probes of various speeds from B. Data adapted from reference [72]. (F) Eye velocity responses after 100 learning trials using different instruction timings, as in C. Data replotted from reference [71]. (G) Complex-spike-linked trial-over-trial change in simple-spike firing for the Purkinje cell shown in D. Grey shaded region highlights the well-timed simple-spike depression due to the occurrence of a complex spike following the instruction from D. (H) Model predicted change in simple-spike responses to the paradigm of B. (I) Model predicted change in

simple-spike firing following a single learning trial with each of the three instruction timings from the experimental paradigm of C. Shaded regions in all panels denotes mean ± SEM across neurons or behavioral learning paradigm replicates.

The model also reproduces the impact of the timing of the instructive stimulus[71] on the magnitude and duration of the learned behavioral response. To study learned timing, we repeated several hundred learning trials with a fixed interval of 150, 250, or 500 ms between the onset of target motion and the time of the instruction[71] (Figure 6C). In the data, longer intervals yielded smaller and temporally-broader behavioral learned responses (Figure 6F). We simulated learned timing in our model by reducing the connection weights to Purkinje cells from parallel fibers that were active at one of the three intervals post-target motion onset: 150 ms, 250 ms, and 500 ms (colored arrows in Figure 6I). The resulting changes in modeled Purkinje cell simple-spike responses (Figure 6I) were consistent with the timing-dependent characteristics observed in learned eye movement responses (Figure 6F). The dependence of modeled learned response on the timing of the instructive stimulus is due to positive correlation between peak and duration time constants for granule layer units. Granule cell axons subject to plasticity due to an early complex spike have smaller response widths than do the granule cells active later in the pursuit trial. Comparable temporal basis sets provide a neural mechanism for the learned timing in eye blink conditioning[66].

Links between complex spikes and Purkinje cell directional and temporal tuning

By genetically altering the projections from the inferior olive to the flocculus in mice, an elegant study demonstrated that the directional modulation of Purkinje cell simple spikes is dictated by the directional preference of climbing fibers[76]. With this in mind, we wondered whether the simple-spike responses measured in floccular Purkinje cells in the monkey during pursuit may have emerged as a consequence of the tuning of climbing fibers and whether climbing fibers might control direction tuning, temporal dynamics, or both. Therefore, we investigated the potential relationship between complex-spike responses and Purkinje cell simple-spike activity, now during baseline pursuit rather than in relation to pursuit learning.

We observed an anti-correlation between the directional preferences of simple spikes and complex spikes, both in absolute preferred direction[22,46,47,77,78] (Figure 7A) and, for both the preferred and non-preferred directions, in the magnitude of the simple-spike response during pursuit and the probability of complex spikes (Figure 7B, $R^2 = 0.16$, $t(196) = -6.2$, $p < 10^{-9}$). Here, we aligned Purkinje cell simple- and complex-spike responses to the preferred direction of complex spikes (CS-on), defined as the direction of pursuit that elicited the maximum increase in complex-spike responses relative to baseline from pursuit onset to offset. In the CS-on direction, higher probabilities of complex spikes were associated with larger inhibitions of simple-spike firing during pursuit. In the CS-off direction, lower probabilities of complex spikes were associated with larger increases in simple-spike firing. Thus, complex-spike firing could dictate simple-spike direction tuning.

The relationship between the temporal response properties of simple and complex spikes is less obvious and emerged clearly only when we considered both the modulation provided by molecular layer interneurons and the properties of the granule unit temporal basis set. Alone, the temporal dynamics of complex-spike firing (Figure 7C) aligns poorly with that of simple-spike

firing during pursuit. In the CS-on direction, the inverse of complex-spike firing starts earlier, rises faster, and decays more quickly compared to simple-spike firing (black vs. red traces, bottom graph of Figure 7D). In the CS-off direction, the inverse of complex-spike firing shows both an early transient increase and a positive spike after the end of target motion (cyan trace, Figure 7C), neither of which has a correlate in simple-spike firing (compare black vs. cyan traces, top graph Figure 7D). Thus, a linear model linking the timeseries of complex-spike activity to modulation of Purkinje cell simple-spike responses across our population showed limited predictive power but consistently negative correlations (Figure 7E, $r = -0.14 \pm 0.03$; $t(99) = -5.4$, $p < 10^{-5}$ in the CS-on direction and $r = -0.24 \pm 0.02$; $t(99) = -11.3$, $p < 10^{-18}$ in the CS-off direction).

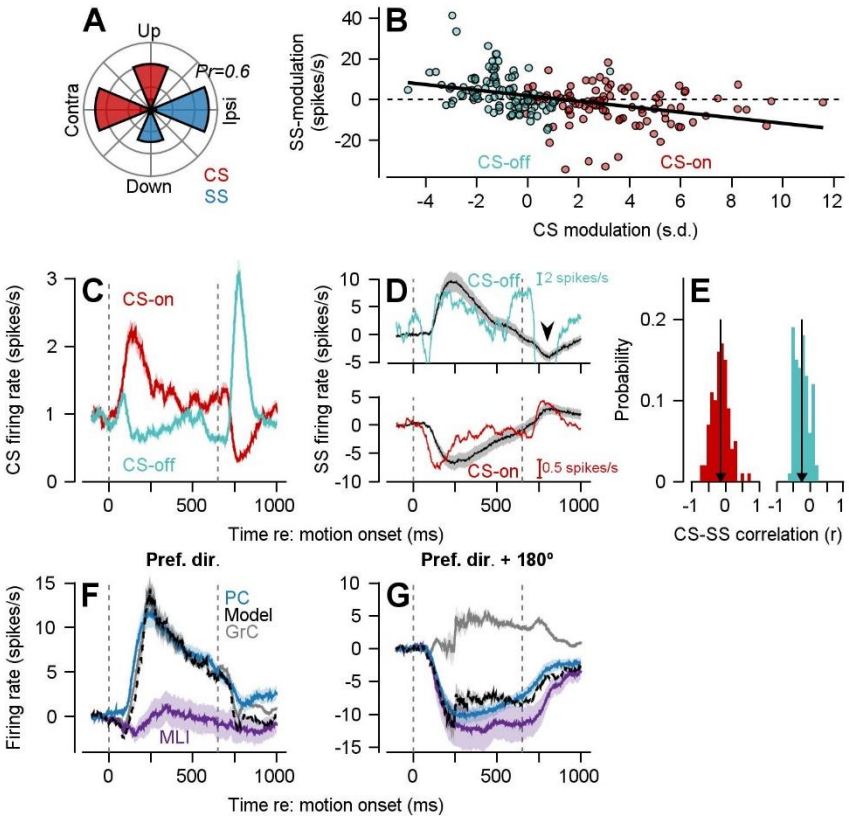

Figure 7. Evidence that complex spike mediated plasticity may create and/or maintain aspects of the temporal and directional dynamics of Purkinje cell firing. (A) Distribution of preferred pursuit directions for Purkinje cell simple spikes (blue) and their associated complex spikes (red). (B) Relationship between simple-spike modulation during pursuit in the CS-on (preferred CS) and CS-off (preferred CS + 180°) directions as a function of complex-spike modulation across all Purkinje cells. Black line denotes best linear fit. (C) Red and teal traces show firing rate of complex spikes in the preferred and anti-preferred complex-spike direction, smoothed using a 50 ms boxcar filter and averaged across Purkinje cells. (D) Comparison of Purkinje cell simple-spike modulation versus scaled and inverted complex-spike modulation in the CS-off (top) and CS-on (bottom) directions. Arrowhead denotes a region of particularly poor fit between simple-spike and complex-spike modulation. Colored and black traces show complex-spike and simple-spike firing. (E) Histogram of Pearson's correlation coefficient between the timeseries of complex-spike activity shown in C and simple-spike activity

shown in D in the CS-on and CS-off directions. Vertical lines denote the population mean. (F) Predicted simple-spike firing from a model that incorporates the granule unit temporal basis set, complex-spike firing linked plasticity, and molecular layer interneuron inputs to create the mean Purkinje cell firing rates (blue trace) in the preferred simple-spike direction. (G) Same as F for the anti-preferred simple-spike direction. Note that the simple spike traces differ between panels D and F-G due to the alignment direction (CS-on versus simple-spike preferred direction). Shaded regions in all panels denotes mean ± SEM across neurons.

We found a tight link between the temporal properties of simple- and complex-spike firing when we used the successful circuit model of Figure 5 to adapt the parallel fiber to Purkinje cell synapses on the basis the complex-spike responses of our population of Purkinje cells during pursuit. To obtain the fits in Figure 7F and G, we decreased the weights of individual granule cell inputs to Purkinje cells based on the product of each granule unit's response during pursuit and the mean complex-spike response measured in the same direction. From the excitatory input of granule cells, we subtracted the weighted input of molecular layer interneurons. Thus, the fits shown in Figure 7F and G had two unknowns: the magnitude of parallel fiber synaptic depression due to a complex spike and the weight of inputs from molecular layer interneurons.

The excellent match between observed and predicted mean simple-spike firing (Figure 7F and G) implies that climbing-fiber inputs coupled with granule layer activity could contribute to the temporal dynamics of cerebellar output. As before, the success of the predictions depends on 1) the temporal basis set in the granule unit population, 2) refinement of the strength of inputs from the granule layer basis set by complex-spike mediated plasticity, and 3) direction-tuned inhibition from molecular layer interneurons to create decreases in simple-spike firing because plasticity of the inputs from spontaneously-silent granule cell inputs cannot drive Purkinje cell responses below baseline.

Discussion

The brain generates behavior through the operation of neural circuits that transform the dynamics of their inputs[79–81] to generate outputs that facilitate appropriate downstream neural processing[82]. As neural circuits are comprised of multiple neuron types[9,83–90], a full understanding of the nature of the transformations mandates a multi-level understanding of the "parts list" of a brain area via both identification of the characteristic neuron types and description of their functional characteristics during behavior. Here, we demonstrate that discrete neuron types within a well-defined neural circuit have distinct functional roles that allow the circuit to perform fundamental transformations of population dynamics to facilitate behavioral output. We also show how complex-spike mediated plasticity complements the distinct neuron-type computations to facilitate the transformations between mossy fiber inputs and eventual Purkinje cell outputs.

Computation by a neural circuit

To understand how a specific circuit computes, we took advantage of new technology that allowed us to identify neuron types in the cerebellar cortex from extracellular recordings[23,28]. The canonical cerebellar circuit, including its intrinsic interneurons, has a conserved architecture across regions and species[7,8] and supports a wide range of behaviors ranging from motor control[15] to communication[91,92] to cognition[93,94]. We recorded the temporal dynamics and directional organization of firing in identified mossy fiber input elements and Purkinje cell

output neurons as well as from intrinsic interneurons: unipolar brush cells, Golgi cells, and molecular layer interneurons, all during a repeatable, quantifiable behavior. Further, we extended the measured responses to identify the necessary properties of the granule cell population, even though we lack the ability to directly record its responses using current technologies.

We suggest that an input-output transformation of temporal dynamics occurs in the cerebellar input layer, potentially via synaptic[66] or granule cell intrinsic[95,96] mechanisms. A directional transformation occurs through circuit mechanisms in the last-order inputs to the output neurons. We confirmed the veracity of the proposed computation through simulation of a circuit model that predicts the responses of each individual Purkinje cell in our sample. Further, the model has multiple emergent properties that agree with known experimental observations[71,72]. We think that we have explained how the cerebellar circuit computes during a specific sensory-motor behavior that requires the region of the cerebellum we studied, namely the floccular complex.

The properties of the temporal basis set provided by the granule unit population were critical for the success of our model. They were: 1) a broad range of temporal filter responses such that granule units temporally tile complete trial; 2) strongly correlated onset and offset time constants resulting in scalar variability[63]; 3) granule unit population activity that increased with increasing pursuit speeds; and 4) consistent timing of individual granule unit activations across pursuit speed to enable generalization of learning across pursuit speeds. These characteristics are consistent with previously proposed "spectral timing" models[39,64,65], and allow us to effectively discount alterative models of granule unit processing[64]. We can make quite specific predictions about granule cell activity because (1) we measure the mossy fiber inputs and (2) we have far more degrees of behavioral freedom than other tasks used to place limits on granule layer processing to-date. For example, cerebellar-dependent eyelid conditioning has a single primary degree of freedom: the timing between the conditioned and unconditioned stimuli. We can vary timing, pursuit speed, and pursuit direction, enabling more strenuous constraints on our model of circuit processing.

Our analysis provides a roadmap for asking whether the cerebellar cortex performs a "universal transform" that aligns with the universal architecture of the cerebellar circuit[97,98]. For example, there is enough data about the responses of neurons in the eye blink area in mice[99] or rabbits[100] to deploy our analysis approach in the eye blink cerebellar region and ask whether a universal transform exists across these two behaviors. More generally, it seems plausible that many or all regions of the cerebellar cortex compute temporal transformations, and the temporal basis set provided by our model granule units could support many different temporal transformations on the time scale of movements. Further, the multiple plasticity mechanisms[101,102] in the cerebellar cortex could adapt the transformations so that they are specialized across behaviors and areas. Meta-plasticity[103] combined with the general nature of the granule unit basis set may provide an even broader ability for the cerebellum to provide region-specific and task-specific circuit transformations.

Our strategy for understanding circuit transformations certainly generalizes to different cerebellar regions and likely to other brain areas. Its key features were: 1) an ability to map extracellular signals to cerebellar neuron types[28] that is validated across species for multiple regions of the cerebellum[23], 2) interpretational context that provides an understanding of how modulation of firing rate in floccular Purkinje cells alters eye movements[104,105], 3) a region that

performs quantifiable temporal and directional transformations of its inputs within a broader brain network that is understood quite thoroughly[106], 4) quantitative behavioral paradigms to characterize the temporal and spatial organization of inputs, outputs, and internal processing, and 5) computational analyses to validate the proposed circuit model.

Specifics of floccular circuit computation

Temporal processing. In the floccular complex, mossy fiber inputs signal eye position with a modest relation to eye velocity[38]; Purkinje cell outputs are related to eye velocity[107] with a modest relation to eye acceleration[35]. Our data imply that the input-output transformation of temporal dynamics occurs in the granule cell layer, potentially via cellular mechanisms at the mossy fiber to granule cell synapse[66]. We localized the transformation to that synapse through the observation that all downstream neurons, including Golgi cells, molecular layer interneurons, and Purkinje cells, show evidence of the full temporal transformation.

Our analyses suggest the role of two kinds of granule layer interneurons, Golgi cells and unipolar brush cells, in distinct aspects of granule cell activity. While we predicted *a priori* that Golgi cells would be instrumental in actively performing temporal transformations, and they may be in other systems, they proved to be only weakly temporally modulated (in relation to eye velocity) in our system. They seem to provide tonic inhibition that may help to create a sparse but temporally diverse[108] representation in the granule cells. Consistent with *in vitro* results[55,56], unipolar brush cells seem to temporally integrate the characteristics of upstream mossy fiber inputs, potentially diversifying the temporal inputs to granule cells[109,110].

An engineer would not perform temporal transformations by creating a temporally-sparse basis set as an intermediate population code that transforms sustained inputs into a temporally-distributed set of transient responses. Yet, biology may use that strategy because it allows a broad range of transformations of temporal dynamics, not just mathematical differentiation, and has the added advantage of affording temporally-specific learning of sensory-motor transformations. Sadly, current extracellular recording technology does not yet allow access to granule cell responses[23]. Thus, our theory, like others, makes a prediction that will need to be tested once technology has evolved.

Directional processing. In the floccular complex, mossy fiber inputs are organized according to the cardinal directions – left, right, up, down – while Purkinje cell outputs prefer “ipsiversive” (towards the side of recording) or downward plus slightly contraversive eye motion[36]. Our analysis of the direction tuning in different neuron types implies that parallel fibers with all direction tunings cross the dendrites of individual Purkinje cells. Multi-directional signals are necessary in the parallel fiber inputs to each Purkinje cell, both to allow cerebellar learning to tune the directionality of Purkinje cells and to enable directional learning in pursuit[72]. Thus, unlike the temporal transformation in the input layer, the directional transformation in the floccular complex appears to occur in the molecular layer.

Molecular layer interneurons play key roles in directional tuning for both baseline pursuit responses and learning. For baseline pursuit responses, they inhibit Purkinje cells; their modulation of firing rate for eye movement in the non-preferred direction drives Purkinje cell simple-spike responses below their spontaneous firing rates. During learning, molecular layer interneurons provide necessary inhibition that allows a directional depression of simple-spike

firing[46,47,73,75] through complex-spike mediated synaptic depression of directional parallel fiber to Purkinje cell synapses. We note that our statements about molecular layer interneurons are limited to the class called "MLI-1", which inhibit Purkinje cells[27] and constitute our entire expert-identified sample. We cannot make any statements about the class called "MLI-2", which inhibit other MLIs but were not included in our sample.

Functional significance of the input-output transformations. Why is it necessary to functionally differentiate the signals of mossy fiber inputs to the floccular complex? We assume that the inputs to the flocculus are limited by the "language" spoken by the brainstem[107], where a "neural integrator" transforms transient inputs into the sustained firing needed to maintain stable eye position at any eccentric position[111]. Thus, the corollary discharge inputs to the floccular complex signal eye position. Yet, the output from the floccular complex must interface in the brainstem with the head-velocity-related inputs from the vestibular system and thus must signal eye velocity. The cerebellum's transformations need to adapt to the fixed and immutable demands of the sensory and motor periphery.

Limitations of the study

We see several limitations. First, our study is correlative in nature and thus constrained by the experimental conditions we tested. However, the ability of our cerebellar circuit model to explain past experimental observations[71,72] suggests that we have obtained a fundamental understanding of how the circuit operates during pursuit. Second, we could not record granule cells and had to rely on computational modeling and intrinsic knowledge of the cerebellar circuitry to place limits on the properties of the granule cell population. Future advances in recording technology may allow us or others to test our predictions of granule layer processing. Finally, we recorded and described the activity of the principal neuron types in the cerebellar circuit but did not characterize several neuron types that remain difficult to identify (i.e., candelabrum, Lugaro, and globular cells). The role of these "Purkinje layer interneurons" awaits technology for recording and identifying them definitively.

Acknowledgements

This study was supported by NIH grants R01-NS112917 (SGL), R01-EY035241 (SGL), and K99-EY030528 (DJH). We thank Stefanie Tokiyama and Bonnie Bowell for assistance with animal care. We also thank Nathan J. Hall and Seth W. Egger for helpful comments and discussions.

Author contributions

DJH and SGL designed all experimental and analysis procedures. DJH performed extracellular recordings in the cerebellar circuit. DJH analyzed the data. DJH and SGL generated the figures and wrote the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

Methods

All experimental procedures were approved in advance by the Duke *Institutional Care and Use Committee* under protocols A085-18-04, A062-21-03, and A016-24-01. All animal care and experimental procedures followed the guidelines outlined in the *NIH Guide for the Care and Use of Laboratory Animals* (1997). Three rhesus macaques (*macaca mulatta*, all male, 10-15 kg) were used in this study. Portions of the dataset analyzed for this manuscript have been included in three previous reports[22,23,28].

General procedures

Prior to behavioral training or neurophysiological recordings, monkeys were instrumented with a head-restraint device, an eye coil, and at least one recording cylinder via separate surgical procedures, all performed with aseptic technique. For each surgical procedure, animals were deeply anaesthetized with isoflurane. Animals received peri- and post-operative analgesia until they had completely recovered. In the first of multiple surgical procedures, head restraint hardware was attached to the animal's skull to allow measurement of the animal's eye position without concomitant head movements. In a later surgery, we sutured a small coil of wire to the sclera of one eye[112], allowing measurement of the animal's eye position using the search coil technique[113]. Following these two surgical procedures, animals were trained to perform smooth pursuit eye movements during discrete trials of target motion in exchange for fluid reward. The experimental paradigm is described in detail below. Once the animal had demonstrated excellent smooth pursuit tracking abilities, as evidenced by minimal intervening saccadic eye movements, we affixed a stainless-steel recording cylinder above a craniectomy, allowing access to the floccular complex of the cerebellum with microelectrodes. The position of this cylinder was 11 mm lateral to the midline, angled at 26° relative to the frontal plane, and pointed towards the interaural axis.

Behavioral task

Each day animals were seated in a dark room, 30 cm in front of a CRT monitor (80 Hz refresh rate, 2304 x 1440 pixels, 480 x 310 mm). Their heads were attached to the chair via the previously implanted restraint hardware. Horizontal and vertical eye position signals were separately digitized at 1,000 Hz using the search coil system. We computed the velocity of the animal's eye movement offline using a 2nd order causal Butterworth low-pass filter with a cutoff frequency of 10 Hz. As we were principally interested in the relationship between floccular neural responses and smooth pursuit eye movements, we removed any saccadic eye movements from ongoing pursuit using an automated procedure[22]. The occurrence of a saccade was identified offline using eye velocity (50 deg/s) and eye acceleration thresholds (1,250 deg/s^2). Onset and offset of each saccade were determined as the time when the animal's eye velocity or acceleration fell below both thresholds for more than 10 ms. In all analyses, we treated eye kinematics during saccades as missing data.

Stimulus presentation was controlled via our lab's custom "Maestro" software. The visual stimulus in all trials was a small (0.5° diameter) black spot shown on a light gray background. In a small subset of trials, animals were required to fixate a stationary target placed at one of nine locations, evenly spaced within a 10° x 10° square. The monkey was required to fixate the dot at each position (within ±1°) continuously for one second in exchange for reward. In the majority of trials, animals tracked smoothly moving targets. Each smooth pursuit trial began by fixating the dot in the center of the screen for a randomly chosen intertrial interval (400-800 ms,

uniformly distributed). After the fixation interval, we used Rashbass' step-ramp paradigm[114]. On each trial, the target moved in the "pursuit direction" with a speed of 10, 20, or 30 deg/s. At the onset of target motion, the target was stepped backwards by 1.5°, 3°, or 4.5° for each of the three target velocities. The backwards step minimizes catch-up saccades caused by the visual latency of the pursuit system[115]. The target moved at constant velocity for 650 ms before stopping at an eccentric position for an additional 350 ms. The majority of pursuit trials were performed in the four cardinal directions. We also performed some trials spaced every 45 degrees, but only for 20 deg/s target speeds. Animals were rewarded for keeping their eyes within an invisible 3° bounding box centered on the pursuit target for the duration of the trial, including initial fixation, pursuit, and eccentric fixation at the end of the trial. Animals were trained extensively on the pursuit task prior to neurophysiological recordings.

Neurophysiology recordings

All recordings were made in the ventral paraflocculus and flocculus, a region we call the floccular complex. These regions have been shown to be crucial for the execution of smooth eye movements[18]. In addition, electrical stimulation of this region of the cerebellum drives smooth movement of the eye towards the side of stimulation[31,32]. We acutely inserted either single tungsten microelectrodes (FHC, 1-2 MΩ) or, more commonly, custom designed Plexon S-Probes into the brain each day through the previously implanted recording cylinder. Plexon S-probes featured 16 tungsten recording contacts, each 7.5 µm in diameter. The 16 contacts were arranged in two columns (50 x 50 µm separation between adjacent contacts). Almost all of our recordings come from the S-Probes.

We identified the floccular complex by its strong activity during smooth pursuit eye movements, presence of infrequent Purkinje cell complex spikes, and the depth relative to the tentorium. Upon arriving in the floccular complex, we waited a minimum of 30 minutes, up to several hours, before beginning neurophysiological recordings. This initial period maximizes the stability of the recording by minimizing drift of neural units across the recording contacts.

Neurophysiological data were recorded using the Plexon Omniplex system. We used analog Butterworth low-pass hardware filters (4^{th} order, 6 kHz cutoff) prior to digitization to minimize any interference from the eye coil system and prevent aliasing. Wideband neural activity on each contact was recorded at 40 kHz, synchronized to behavioral data using high speed TTL pulses, and stored for later offline analysis.

To identify well-isolated single neurons from our wideband voltage recordings, we leveraged the Full Binary Pursuit[116] (FBP) spike sorting package. The FBP sorter is specifically designed to optimally resolve temporally and spatially overlapping action potentials, a frequent occurrence in the cerebellar cortex due to the relatively high baseline firing rates of many cerebellar neurons. Following spike sorting, we manually curated all neural units to ensure high quality single-units. We specifically excluded neural units that showed evidence of contamination by either background noise or other units. We assayed contamination primarily via assessment of refractory period violations, measured as the rate of spikes that violated a presumed 1 ms absolute refractory period. Across all monkeys, we recorded *n=1,152* single units. Across this sample of neurons, mean refractory period violations were 0.6 ± 2.5% (mean ± SD) of all recorded spikes. We measured the maximum peak-to-peak voltage deviation of the mean action potential waveform/template on the channel with the largest spike and compared this amplitude

to the level of background noise measured on the same channel (the primary contact). The mean signal-to-noise ratio, defined as the amplitude of the spike divided by an estimate of noise computed as 1.96-times the standard deviation of background noise, was 5.6 ± 2.9 (mean ± SD). We converted neuron spike trains into firing rates by causally convolving each spike train with a dual exponential kernel. The kernel (τ_{rise} = 0.1 ms and τ_{decay} = 50 ms) mimics the properties of post-synaptic currents[117] while preserving the ability to measure latencies with high precision. Autocorrelograms were computed using previously described techniques[22], and normalized by the bin width (1 ms), placing autocorrelograms in units of spikes/s.

Expert identification of cerebellar neuron types

A detailed discussion of the methodology for identification of cerebellar neuron types in our recordings has been published previously[28]. Briefly, we began by identifying a subset of units as ground-truth Purkinje cells based on their unique extracellular properties. Purkinje cells receive excitatory inputs from parallel fibers as well as strong inputs via climbing fibers from the inferior olive that drive post-synaptic Purkinje cell complex spikes. During complex spikes and for several milliseconds thereafter, Purkinje cells are prevented from firing simple spikes, resulting in a stereotypical complex-spike-induced pause in simple spikes. We included only Purkinje cells that had a complex-spike-induced pause.

We identified molecular layer interneurons by establishing monosynaptic functional inhibitory relationships between these neurons and simultaneously recorded ground-truth Purkinje cells. Therefore, our sample of molecular layer interneurons constitutes principally MLI-1's, classified by others based on their genetic[13] and connection profiles[27]. Golgi cells were identified by their low firing rate responses, characteristic extracellular waveforms[118], and presence within the granule cell layer. Mossy fiber inputs to the cerebellar cortex were identified based on the presence of a negative-after-wave (NAW). The NAW corresponds to the postsynaptic response of granule cells within the cerebellar glomerulus[119]. We note that while a NAW is sufficient to identify mossy fibers, it is possible to record from mossy fiber axons that do not show marked NAWs. To ensure that our sample contained only known mossy fibers, we excluded from analysis putative mossy fibers that lacked a NAW. Unipolar brush cells were identified based on established functional responses recorded by *in vitro* studies showing that unipolar brush cells elongate the timescales of discrete mossy fiber input over 10s to hundreds of milliseconds[55,56]. We considered units to be unipolar brush cells if triggering their firing off simultaneously recorded mossy fiber bursts yielded temporally-elongated responses. We recently validated our expert-labeling approach using a deep-learning classifier trained on ground-truth optogenetically-identified recordings in mice[23].

Principal component analysis

To identify the primary modes of temporal information contained in the cerebellar population during smooth pursuit eye movements, we performed principal component analysis across all expert-identified neurons. As the responsiveness of cerebellar neurons to pursuit varied widely, we wanted to ensure that highly responsive neurons did not bias our estimate of the principal component directions. Therefore, we performed soft-normalization of neuron firing rates[45], measured in their preferred direction, prior to principal component analysis. Soft-normalization squashes modulations less than 5 spikes/s to timeseries near zero and scales modulations of greater than 5 spikes/s to approximately unit magnitude. After soft-normalization, we smoothed the firing rate traces using a boxcar filter with a 50 ms width. To identify the principal

components across speed and direction for mossy fibers and Purkinje cells, we used the same preprocessing procedures separately in each respective population. Then, we concatenated the within-neuron responses to different speeds and directions, resulting in an N x ($S*D*N_t$) matrix, where N represents the number of units in each population, S represents the number of speeds (three: 10, 20, 30 deg/s), D the number of directions (four: ipsiversive, up, contraversive, down), and N_t the number of timepoints in each sample (1,100). As with the complete population, principal components analysis identified the modes of temporal profiles across the neuron dimension in each population.

Regression of neuron firing rates to eye kinematics

To evaluate the contributions of different eye kinematic signals to the firing rates of cerebellar neuron types, we performed linear regression analysis by fitting kinematic models to each neuron's mean time-varying firing rate during smooth pursuit. Depending on whether the firing in the anti-preferred direction hit a floor at zero firing rate, we fit the data either in the neuron's preferred direction or across both its preferred and anti-preferred directions simultaneously. For each neuron, we fit models of the following form:

$$r(t) = w_p e(t + \tau_d) + w_v \dot{e}(t + \tau_d) + w_a \ddot{e}(t + \tau_d) + b \quad (1)$$

In Equation 1, *r(t)* represents the mean firing rate of the neuron across pursuit trials and is expressed as a weighted sum of the mean eye position (e), velocity ($\dot{e}$), and acceleration ($\ddot{e}$) at a future timepoint, ($t + \tau_d$). The parameter τ_d refers to the temporal lead of the firing rate relative to the kinematics. The scalar parameter, b, is a neuron-specific bias. We determined the unknown parameters, $\{w_p, w_v, w_a, b\}$ using least squares for each delay from 0 to 100 ms. The optimal delay was identified by minimizing the mean squared error across the range of delays. To assess the relative contribution of each kinematic variable to the overall firing rate, we computed standardized (β) coefficients by normalizing each weight by the ratio of the standard deviation of the corresponding regressor to the standard deviation of the observed firing rate.

Recurrent neural network model of granule cell transformations

We constructed a recurrent neural network model to investigate features of the transformation between mossy fiber inputs and Purkinje cell outputs. We supplied as inputs to the model single-trial firing rate responses of individual mossy fibers ($n = 86$) during 10, 20, or 30 deg/s smooth pursuit trials in the preferred direction of each mossy fiber. We trained the complete network model to predict the single-trial firing responses of our recorded Purkinje cell population. Purkinje cell responses were in their preferred directions with speeds that were trial-matched to the speed of pursuit supplied as mossy fiber inputs (e.g., mossy fiber inputs corresponding to single-trial responses during 10 deg/s target motion were paired with Purkinje cell outputs also measured during 10 deg/s pursuit). To ease training, we preconditioned the input and output data by soft-normalizing[45] the responses of the mossy fiber and Purkinje cell responses. The soft-normalized mossy fiber inputs were supplied to a pool of RNN units, whose dynamics were governed by:

$$\tau \dot{\boldsymbol{x}}(t) = -\boldsymbol{x}(t) + \sigma(W_i \boldsymbol{u}(t) + W_r \boldsymbol{x}(t) + \boldsymbol{b}) \quad (2)$$

In Equation 2, the activity of each of 25 RNN units is represented by the vector $\boldsymbol{x}$. The parameter τ corresponds to the time constant of network integration (fixed at 10 ms). The matrix W_r

(dimensions 25 x 25) serves as recurrent weights on RNN unit activity from the previous timestep while the matrix W_i (dimensions 25 x 86) weighs incoming mossy fiber inputs, represented by $\boldsymbol{u(t)}$. The vector $\boldsymbol{b}$ serves as a time-invariant bias. The combination of inputs from upstream mossy fiber activity, recurrent inputs, and the bias parameter was passed through a non-linear activation function, σ, chosen to be the sigmoid function. The reservoir of RNN units served as the temporal bases for the subsequent "granule unit" layer. Each granule unit received a weighted set of inputs from the 25 RNN units. Granule unit inputs were passed through a sigmoid activation function, constraining the activity of individual granule units between zero and one. Finally, a set of fully connected weights mapped the $n = 100$ granule units onto each output Purkinje cell. Unlike previous layers, we constrained the weight matrix between granule units and Purkinje cells to be strictly positive. The complete network was simultaneously trained using the Adam[120] optimizer with cosine annealing with warm-restarts[121]. We used drop-out layers during training with 50% probability of dropout before and after the granule unit layer to avoid overfitting and ensure that the dynamics necessary to represent Purkinje cell firing rates were distributed throughout the granule layer units. We used early termination to stop training when the cross-validated error from a withheld 10% of the single trial data failed to decrease for more than 10 training epochs.

We evaluated performance of the network by supplying as input the trial-averaged activity (rather than single-trial activity) of mossy fibers and measuring the network predicted Purkinje cell responses. We compared the predicted responses to the mean measured, soft-normalized, Purkinje cell responses across target speeds. We also examined the nature of the computation performed by the complete network. To do so, we supplied a step-response of mossy fiber activity as input. Prior to $t=0$ ms, the input to the network corresponded to the soft-normalized baseline activity of each mossy fiber, as measured before target motion onset. At $t=0$ ms, we stepped the response of each mossy fiber to the soft-normalized response measured at the termination of target motion across pursuit in the 20 deg/s condition. We then interrogated the responses of the granule units across time from before to the end of the step change in input.

A generative model of granule cell dynamics

Our goal was to establish a simplified mathematical description of granule cell processing that recapitulated the primary results established by our recurrent neural network model. Namely, the granule units from the RNN model showed 1) a broad range of integration time constants (time to peak) following a step input, 2) neurons with a later peak showed broader temporal responses, 3) larger amplitude responses across the granule population for faster pursuit speeds, and 4) consistent peak timing across pursuit speeds in individual granule units. While multiple models are likely capable of showing these primary features, our implementation focused on a non-mechanistic description whose parameter values can be easily interpreted relative to their functional outcomes. We emphasize that the generative model is not meant to be mechanistically biomimetic, merely operational to explore the features of successful temporal basis sets.

We began by establishing synaptic input to each granule unit:

$$u_n(t) = \frac{\sum_{i=1}^{4} w_i MF_i(t)}{\sum_{i=1}^{4} w_i} \quad (3)$$

where, u_n represents the total synaptic input to the granule unit. The four mossy fibers were chosen randomly from our sample irrespective of their preferred directions and individual weights of mossy fibers, w_i, were chosen from a uniform distribution between 0 and 1. The sum of the weights was normalized to unity magnitude. The dynamics of each granule unit are described by a set of continuous-time differential equations:

$$\dot{\boldsymbol{x_n}}(t) = \begin{bmatrix} \alpha & 1 \\ -k^2 & -2k \end{bmatrix} \boldsymbol{x_n}(t) + \begin{bmatrix} 0 \\ k^2 \end{bmatrix} (u_n(t) - b_n)$$
$$r_n(t) = \left[[0 \quad 1] \boldsymbol{x_n}(t) \right]_+ \tag{4}$$

Equation 4 represents a generalization of a differentiating dynamical system with two states, $\boldsymbol{x_n}$. The input to the system is the weighted mossy fiber inputs from Equation 3, u_n, less a unit specific bias, b_n. Each unit's bias was uniformly distributed in the range [50, 100] spikes/s. The relatively large bias serves to sparsify the granule unit responses; alternative distributions for the bias parameter do not impact qualitatively our findings. The primary free parameter is *k*, which controls the speed of differentiation. Large values of *k* result in rapid differentiation while small values of *k* result in slower and thus smoother differentiation. We chose *k* to be uniformly distributed in the range [0, 50]. The final parameter, α, controls the amount of low-pass information retained in the output, where increasingly positive values of α result in incomplete differentiation. We chose α from a uniform distribution with range [0, 0.1]. The output of the system, r_n (corresponding to the second state, x_2) is constrained to be greater than zero (half-wave rectification), indicated by $]_+$. Thus, the dynamical system described in Equation 4 is a tunable differentiator whose output is half-wave rectified. We emphasize that the functional properties of the temporal basis set documented in Figures 4E and F are the crucial components of the granule layer - multiple model formulations would yield comparable results. Indeed, our proposed generative model represents a specific implementation of a spectral timing model for granule cell activity, a framework that has been suggested in various forms for eyelid conditioning[64,65]. For all simulations, we converted the continuous state representation in Equation 4 to its discrete-time equivalent using the forward Euler method.

Cerebellar circuit model for predicting downstream neural responses

Using the generative model of granule cell rate responses (Equations 3-4), we asked whether the resulting granule unit population could be used to replicate the measured firing rate responses of neural units downstream of granule cells: molecular layer interneurons and Purkinje cells. We constructed a population of 1,000 simulated granule units using the procedures outlined in Equations 3-4. For each of the granule units, we computed the predicted firing rate responses to pursuit of 20 deg/s in the cardinal directions. We started by modeling the responses of molecular layer interneurons. Across all directions simultaneously, we found a set of weights that minimized the cost function detailed in Equation 5.

$$J_n = (\boldsymbol{y_n^{MLI}} - R^T \boldsymbol{w_n})^T (\boldsymbol{y_n^{MLI}} - R^T \boldsymbol{w_n}) + \lambda \boldsymbol{w_n^T} \mathrm{diag}\left[\frac{1}{\sigma^2(R)}\right] \boldsymbol{w_n} \tag{5}$$

In Equation 5, $\boldsymbol{y_n^{MLI}}$ is a vector of the *n*-th molecular layer interneuron responses across the cardinal directions (e.g., length N_t*D), *R* is a 1,001 x (N_t*D) design matrix of mean-subtracted simulated granule cell responses across the same directions augmented with a row of ones. As

our population of molecular layer interneurons was relatively small, we augmented our measured population by randomly averaging sets of 5 recorded molecular layer interneurons, with replacement. The cost function in Equation 5 is minimized to find the unknown weights of the granule cell responses, $\boldsymbol{w_n}$ for each molecular layer interneuron. As the granule cell inputs to molecular layer interneurons are exclusively excitatory, we subjected Equation 5 to the constraint $\forall \boldsymbol{w_n} \geq 0$. The parameter λ is a regularization parameter (e.g., Tokhonov or ridge regression parameter) which penalizes large weights. The ridge penalty term is standardized by the standard deviation of each regression parameter in the design matrix (rows in *R*), with the diagonal values of non-varying regressors in *R* set to zero. We found the optimal value of λ using cross-validation by leaving out 50% of the data across time and direction and evaluating the ability of the model to generalize to the withheld data. We evaluated the fits to the molecular layer interneurons using Pearson's correlation coefficients.

After fitting the augmented molecular layer interneuron population, we evaluated the ability to fit our measured Purkinje cell responses across directions. We used the same form of the cost function in Equation 5 except that the design matrix included rows corresponding to the granule units as well as our fitted and baseline-subtracted molecular layer interneuron responses. The rows in the design matrix corresponding to fitted molecular layer interneuron responses were multiplied by -1.0 and directionally permuted so that interneuron preferred directions were opposite Purkinje cell preferred directions. Thus, our goal was to find a set of non-negative weights that described Purkinje cell firing across our population of measured Purkinje cells given upstream firing of granule cells and inhibition by molecular layer interneurons. Just as our procedure to fit the molecular layer interneuron population, we scaled the ridge parameter by the standard deviation of each regressor and found the optimal value of λ using cross-validation.

Evaluation of emergent properties of the cerebellar circuit model

We tested whether our simplified model of cerebellar function could account for previous pursuit learning results. Previous results suggest that short-term pursuit learning is driven by complex-spike mediated plasticity of the parallel fiber to Purkinje cell synapse. A complex spike on a single learning trial is linked to a well-timed depression of Purkinje cell simple-spike firing on the next trial[46,47,73,75,122]. To simulate single-trial depression in our cerebellar circuit model, we modified components of the previously found weights between granule cells and Purkinje cells computed in Equation 5. The change in each granule cell's weight, $\Delta\mathbf{w_n}$, is:

$$\Delta \boldsymbol{w_n} = -k \frac{\boldsymbol{r_n}(\text{Instr.}+75\text{ms})}{\max\left[\boldsymbol{r_n}(\text{Instr.}+75\text{ms})\right]} \tag{6}$$

In Equation 6, $\boldsymbol{r_n}$ is a 1,000-element vector of granule cell activity, relative to baseline, measured at the time of the instruction with an assumed 75 ms visual latency. The scalar *k* is an arbitrary constant that scales the magnitude of Purkinje cell learning. We assayed learning by measuring the simulated change in Purkinje cell simple spikes for the original ($R^T\mathbf{w_n}$) versus the modified weight vector ($R^T[\mathbf{w_n} + \Delta\mathbf{w_n}]$). In one set of simulations, we tested generalization of learning by changing the statistics of the mossy fiber inputs on the test trial after each learning trial, thereby assessing learning for a range of pursuit speeds in the test trial. In a second set of simulations, we changed the timing of the instructive stimulus relative to the onset of pursuit and asked how that altered the Purkinje cell firing on the subsequent test trial.

Statistical analysis

We used the Julia package HypothesisTests for common statistical calculations, including t-tests for correlation and independent samples t-tests. All statistical tests were two-sided and we report exact p-values where possible. To perform permutation tests, we randomly sampled with replacement from two comparison populations under the null hypothesis. We performed 10,000 random permutations for each test unless otherwise noted.

Data availability

All data analyzed for this study have been deposited into the Open Science Framework database. Reasonable requests for additional data or analyses can be made to the corresponding author.

References

1. Vyas, S., Golub, M. D., Sussillo, D. & Shenoy, K. V. Computation Through Neural Population Dynamics. *Annu. Rev. Neurosci.* **43**, 249–275 (2020).
2. Shenoy, K. V. & Kao, J. C. Measurement, manipulation and modeling of brain-wide neural population dynamics. *Nat. Commun.* **12**, 633 (2021).
3. Khilkevich, A. *et al.* Brain-wide dynamics linking sensation to action during decision-making. *Nature* **634**, 890–900 (2024).
4. Runyan, C. A., Piasini, E., Panzeri, S. & Harvey, C. D. Distinct timescales of population coding across cortex. *Nature* **548**, 92–96 (2017).
5. Gold, J. I. & Shadlen, M. N. The Neural Basis of Decision Making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
6. Inagaki, H. K. *et al.* Neural Algorithms and Circuits for Motor Planning. *Annu. Rev. Neurosci.* **45**, 249–271 (2022).
7. Striedter, G. F., Bullock, T. H., Preuss, T. M., Rubenstein, J. & Krubitzer, L. A. *Evolution of Nervous Systems*. (Academic Press, 2016).
8. Jacobs, B. *et al.* Comparative neuronal morphology of the cerebellar cortex in afrotherians, carnivores, cetartiodactyls, and primates. *Front. Neuroanat.* **8**, (2014).
9. Ramón y Cajal, S. *Histologie Du Système Nerveux de l'homme & Des Vertébrés*. (Maloine, Paris, 1909). doi:10.5962/bhl.title.48637.
10. Hull, C. & Regehr, W. G. The Cerebellar Cortex. *Annu. Rev. Neurosci.* **45**, 151–175 (2022).
11. De Zeeuw, C. I., Lisberger, S. G. & Raymond, J. L. Diversity and dynamism in the cerebellum. *Nat. Neurosci.* **24**, 160–167 (2021).
12. Eccles, J. C., Ito, M. & Szentágothai, J. *The Cerebellum as a Neuronal Machine*. (Springer, Berlin, Heidelberg, 1967). doi:10.1007/978-3-662-13147-3.
13. Kozareva, V. *et al.* A transcriptomic atlas of mouse cerebellar cortex comprehensively defines cell types. *Nature* **598**, 214–219 (2021).
14. Osorno, T. *et al.* Candelabrum cells are ubiquitous cerebellar cortex interneurons with specialized circuit properties. *Nat. Neurosci.* **25**, 702–713 (2022).
15. Holmes, G. The Symptoms of Acute Cerebellar Injuries due to Gunshot Injuries. *Brain* **40**, 461–535 (1917).
16. Barash, S. *et al.* Saccadic Dysmetria and Adaptation after Lesions of the Cerebellar Cortex. *J. Neurosci.* **19**, 10931–10939 (1999).
17. Ritchie, L. Effects of cerebellar lesions on saccadic eye movements. *J. Neurophysiol.* **39**, 1246–1256 (1976).
18. Rambold, H., Churchland, A., Selig, Y., Jasmin, L. & Lisberger, S. G. Partial Ablations of the Flocculus and Ventral Paraflocculus in Monkeys Cause Linked Deficits in Smooth Pursuit Eye Movements and Adaptive Modification of the VOR. *J. Neurophysiol.* **87**, 912–924 (2002).
19. Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J. & Thach, W. T. Throwing while looking through prisms. *Brain* **119**, 1183–1198 (1996).
20. Steinmetz, N. A. *et al.* Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. *Science* **372**, eabf4588 (2021).
21. Jun, J. J. *et al.* Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017).
22. Herzfeld, D. J., Joshua, M. & Lisberger, S. G. Rate versus synchrony codes for cerebellar control of motor behavior. *Neuron* **111**, 2448-2460.e6 (2023).

23. Beau, M. *et al.* A deep-learning strategy to identify cell types across species from high-density extracellular recordings. 2024.01.30.577845 Preprint at https://doi.org/10.1101/2024.01.30.577845 (2024).
24. Kostadinov, D., Beau, M., Blanco-Pozo, M. & Häusser, M. Predictive and reactive reward signals conveyed by climbing fiber inputs to cerebellar Purkinje cells. *Nat. Neurosci.* **22**, 950–962 (2019).
25. Sedaghat-Nejad, E., Pi, J. S., Hage, P., Fakharian, M. A. & Shadmehr, R. Synchronous spiking of cerebellar Purkinje cells during control of movements. *Proc. Natl. Acad. Sci.* **119**, e2118954119 (2022).
26. Pi, J. S. *et al.* The olivary input to the cerebellum dissociates sensory events from movement plans. *Proc. Natl. Acad. Sci.* **121**, e2318849121 (2024).
27. Lackey, E. P. *et al.* Specialized connectivity of molecular layer interneuron subtypes leads to disinhibition and synchronous inhibition of cerebellar Purkinje cells. *Neuron* **0**, (2024).
28. Herzfeld, D. J., Hall, N. J. & Lisberger, S. G. Strategies to decipher neuron identity from extracellular recordings in the cerebellum of behaving non-human primates. 2025.01.29.634860 Preprint at https://doi.org/10.1101/2025.01.29.634860 (2025).
29. Belton, T. & McCrea, R. A. Role of the cerebellar flocculus region in the coordination of eye and head movements during gaze pursuit. *J. Neurophysiol.* **84**, 1614–1626 (2000).
30. Highstein, S. M. The organization of the vestibulo-oculomotor and trochlear reflex pathways in the rabbit. *Exp. Brain Res.* **17**, 285–300 (1973).
31. Lisberger, S. G., Pavelko, T. A. & Broussard, D. M. Responses during eye movements of brain stem neurons that receive monosynaptic inhibition from the flocculus and ventral paraflocculus in monkeys. *J. Neurophysiol.* **72**, 909–927 (1994).
32. Joshua, M., Medina, J. F. & Lisberger, S. G. Diversity of Neural Responses in the Brainstem during Smooth Pursuit Eye Movements Constrains the Circuit Mechanisms of Neural Integration. *J. Neurosci.* **33**, 6633–6647 (2013).
33. Bell, C. C. & Grimm, R. J. Discharge properties of Purkinje cells recorded on single and double microelectrodes. *J. Neurophysiol.* **32**, 1044–1055 (1969).
34. Bloedel, J. R. & Roberts, W. J. Action of climbing fibers in cerebellar cortex of the cat. *J. Neurophysiol.* **34**, 17–31 (1971).
35. Lisberger, S. G. & Fuchs, A. F. Role of primate flocculus during rapid behavioral modification of vestibuloocular reflex. I. Purkinje cell activity during visually guided horizontal smooth-pursuit eye movements and passive head rotation. *J. Neurophysiol.* **41**, 733–763 (1978).
36. Stone, L. S. & Lisberger, S. G. Visual responses of Purkinje cells in the cerebellar flocculus during smooth-pursuit eye movements in monkeys. I. Simple spikes. *J. Neurophysiol.* **63**, 1241–1261 (1990).
37. Krauzlis, R. J. & Lisberger, S. G. Directional organization of eye movement and visual signals in the floccular lobe of the monkey cerebellum. *Exp. Brain Res.* **109**, 289–302 (1996).
38. Lisberger, S. G. & Fuchs, A. F. 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.* **41**, 764–777 (1978).
39. Buonomano, D. V. & Mauk, M. D. Neural Network Model of the Cerebellum: Temporal Discrimination and the Timing of Motor Responses. *Neural Comput.* **6**, 38–55 (1994).

40. Medina, J. F., Garcia, K. S., Nores, W. L., Taylor, N. M. & Mauk, M. D. Timing Mechanisms in the Cerebellum: Testing Predictions of a Large-Scale Computer Simulation. *J. Neurosci.* **20**, 5516–5525 (2000).
41. Yamazaki, T. & Tanaka, S. Computational Models of Timing Mechanisms in the Cerebellar Granular Layer. *Cerebellum Lond. Engl.* **8**, 423–432 (2009).
42. Gilmer, J. I. *et al.* An emergent temporal basis set robustly supports cerebellar time-series learning. *J. Neurophysiol.* **129**, 159–176 (2023).
43. Marr, D. A theory of cerebellar cortex. *J. Physiol.* **202**, 437-470.1 (1969).
44. Albus, J. S., Branch, D. T., Donald, C. & Perkel, H. *A Theory of Cerebellar Function*. (1971).
45. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
46. Medina, J. F. & Lisberger, S. G. Links from complex spikes to local plasticity and motor learning in the cerebellum of awake-behaving monkeys. *Nat. Neurosci.* **11**, 1185–1192 (2008).
47. Yang, Y. & Lisberger, S. G. Interaction of plasticity and circuit organization during the acquisition of cerebellum-dependent motor learning. *eLife* **2**, e01574 (2013).
48. Fleming, E. A., Field, G. D., Tadross, M. R. & Hull, C. Local synaptic inhibition mediates cerebellar granule cell pattern separation and enables learned sensorimotor associations. *Nat. Neurosci.* 1–13 (2024) doi:10.1038/s41593-023-01565-4.
49. D'Angelo, E. *et al.* The cerebellar Golgi cell and spatiotemporal organization of granular layer activity. *Front. Neural Circuits* **7**, 93 (2013).
50. D'Angelo, E. The Critical Role of Golgi Cells in Regulating Spatio-Temporal Integration and Plasticity at the Cerebellum Input Stage. *Front. Neurosci.* **2**, 35–46 (2008).
51. Holt, G. R., Softky, W. R., Koch, C. & Douglas, R. J. Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *J. Neurophysiol.* **75**, 1806–1814 (1996).
52. Kim, J. & Augustine, G. J. Molecular Layer Interneurons: Key Elements of Cerebellar Network Computation and Behavior. *Neuroscience* **462**, 22–35 (2021).
53. Mugnaini, E., Sekerková, G. & Martina, M. The unipolar brush cell: A remarkable neuron finally receiving deserved attention. *Brain Res. Rev.* **66**, 220–245 (2011).
54. Balmer, T. S. & Trussell, L. O. Selective targeting of unipolar brush cell subtypes by cerebellar mossy fibers. *eLife* **8**, e44964 (2019).
55. Guo, C., Huson, V., Macosko, E. Z. & Regehr, W. G. Graded heterogeneity of metabotropic signaling underlies a continuum of cell-intrinsic temporal responses in unipolar brush cells. *Nat. Commun.* **12**, 5491 (2021).
56. Huson, V. & Regehr, W. G. Realistic mossy fiber input patterns to unipolar brush cells evoke a continuum of temporal responses comprised of components mediated by different glutamate receptors. *eLife* **13**, (2024).
57. Barmack, N. H. & Yakhnitsa, V. Functions of interneurons in mouse cerebellum. *J. Neurosci.* **28**, 1140–1152 (2008).
58. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
59. Rajalingham, R., Piccato, A. & Jazayeri, M. Recurrent neural networks with explicit representation of dynamic latent variables can mimic behavioral patterns in a physical inference task. *Nat. Commun.* **13**, 5865 (2022).

60. Sussillo, D., Churchland, M. M., Kaufman, M. T. & Shenoy, K. V. A neural network that finds a naturalistic solution for the production of muscle activity. *Nat. Neurosci.* **18**, 1025–1033 (2015).
61. Palkovits, M., Magyar, P. & Szentágothai, J. Quantitative histological analysis of the cerebellar cortex in the cat: II. Cell numbers and densities in the granular layer. *Brain Res.* **32**, 15–30 (1971).
62. Jakab, R. L. & Hámori, J. Quantitative morphology and synaptology of cerebellar glomeruli in the rat. *Anat. Embryol. (Berl.)* **179**, 81–88 (1988).
63. Gibbon, J. Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* **84**, 279–325 (1977).
64. Medina, J. F. & Mauk, M. D. Computer simulation of cerebellar information processing. *Nat. Neurosci.* **3**, 1205–1211 (2000).
65. Bullock, D., Fiala, J. C. & Grossberg, S. A neural model of timed response learning in the cerebellum. *Neural Netw.* **7**, 1101–1114 (1994).
66. Barri, A., Wiechert, M. T., Jazayeri, M. & DiGregorio, D. A. Synaptic basis of a sub-second representation of time in a neural circuit model. *Nat. Commun.* **13**, 7902 (2022).
67. Ahmadian, Y. & Miller, K. D. What is the dynamical regime of cerebral cortex? *Neuron* **109**, 3373–3391 (2021).
68. Xue, M., Atallah, B. V. & Scanziani, M. Equalizing excitation–inhibition ratios across visual cortical neurons. *Nature* **511**, 596–600 (2014).
69. Shu, Y., Hasenstaub, A. & McCormick, D. A. Turning on and off recurrent balanced cortical activity. *Nature* **423**, 288–293 (2003).
70. Wehr, M. & Zador, A. M. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* **426**, 442–446 (2003).
71. Medina, J. F., Carey, M. R. & Lisberger, S. G. The Representation of Time for Motor Learning. *Neuron* **45**, 157–167 (2005).
72. Herzfeld, D. J., Hall, N. J., Tringides, M. & Lisberger, S. G. Principles of operation of a cerebellar learning circuit. *eLife* **9**, e55217 (2020).
73. Yang, Y. & Lisberger, S. G. Purkinje-cell plasticity and cerebellar motor learning are graded by complex-spike duration. *Nature* **510**, 529–532 (2014).
74. Hall, N. J., Yang, Y. & Lisberger, S. G. Multiple components in direction learning in smooth pursuit eye movements of monkeys. *J. Neurophysiol.* (2018) doi:10.1152/jn.00261.2018.
75. Yang, Y. & Lisberger, S. G. Role of Plasticity at Different Sites across the Time Course of Cerebellar Motor Learning. *J. Neurosci.* **34**, 7077–7090 (2014).
76. Badura, A. *et al.* Climbing Fiber Input Shapes Reciprocity of Purkinje Cell Firing. *Neuron* **78**, 700–713 (2013).
77. Stone, L. S. & Lisberger, S. G. Visual responses of Purkinje cells in the cerebellar flocculus during smooth-pursuit eye movements in monkeys. II. Complex spikes. *J. Neurophysiol.* **63**, 1262–1275 (1990).
78. Larry, N., Yarkoni, M., Lixenberg, A. & Joshua, M. Cerebellar climbing fibers encode expected reward size. *eLife* **8**, e46870 (2019).
79. Hopfield, J. J. Transforming neural computations and representing time. *Proc. Natl. Acad. Sci.* **93**, 15440–15444 (1996).
80. Hopfield, J. J. & Tank, D. W. Computing with Neural Circuits: A Model. *Science* **233**, 625–633 (1986).
81. Luo, L. Architectures of neuronal circuits. *Science* **373**, eabg7285 (2021).

82. Mathis, M. W., Perez Rotondo, A., Chang, E. F., Tolias, A. S. & Mathis, A. Decoding the brain: From neural representations to mechanistic models. *Cell* **187**, 5814–5832 (2024).
83. Luo, L., Callaway, E. M. & Svoboda, K. Genetic Dissection of Neural Circuits. *Neuron* **57**, 634–660 (2008).
84. Fishell, G. & Heintz, N. The Neuron Identity Problem: Form Meets Function. *Neuron* **80**, 602–612 (2013).
85. Ecker, J. R. *et al.* The BRAIN Initiative Cell Census Consortium: Lessons Learned toward Generating a Comprehensive Brain Cell Atlas. *Neuron* **96**, 542–557 (2017).
86. Zeng, H. & Sanes, J. R. Neuronal cell-type classification: challenges, opportunities and the path forward. *Nat. Rev. Neurosci.* **18**, 530–546 (2017).
87. Migliore, M. & Shepherd, G. M. An integrated approach to classifying neuronal phenotypes. *Nat. Rev. Neurosci.* **6**, 810–818 (2005).
88. Gouwens, N. W. *et al.* Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nat. Neurosci.* **22**, 1182–1195 (2019).
89. Poulin, J.-F., Tasic, B., Hjerling-Leffler, J., Trimarchi, J. M. & Awatramani, R. Disentangling neural cell diversity using single-cell transcriptomics. *Nat. Neurosci.* **19**, 1131–1141 (2016).
90. Josh Huang, Z. & Zeng, H. Genetic Approaches to Neural Circuits in the Mouse. *Annu. Rev. Neurosci.* **36**, 183–215 (2013).
91. Gordon, N. Speech, language, and the cerebellum. *Eur. J. Disord. Commun. J. Coll. Speech Lang. Ther. Lond.* **31**, 359–367 (1996).
92. Schmahmann, J. D. & Sherman, J. C. The cerebellar cognitive affective syndrome. *Brain* **121**, 561–579 (1998).
93. Van Overwalle, F. Social and emotional learning in the cerebellum. *Nat. Rev. Neurosci.* **25**, 776–791 (2024).
94. Schmahmann, J. D. The cerebellum and cognition. *Neurosci. Lett.* **688**, 62–75 (2019).
95. Straub, I. *et al.* Gradients in the mammalian cerebellar cortex enable Fourier-like transformation and improve storing capacity. *eLife* **9**, e51771 (2020).
96. Masoli, S., Tognolina, M., Laforenza, U., Moccia, F. & D'Angelo, E. Parameter tuning differentiates granule cell subtypes enriching transmission properties at the cerebellum input stage. *Commun. Biol.* **3**, 1–12 (2020).
97. Diedrichsen, J., King, M., Hernandez-Castillo, C., Sereno, M. & Ivry, R. B. Universal Transform or Multiple Functionality? Understanding the Contribution of the Human Cerebellum across Task Domains. *Neuron* **102**, 918–928 (2019).
98. Schmahmann, J. D. From movement to thought: Anatomic substrates of the cerebellar contribution to cognitive processing. *Hum. Brain Mapp.* **4**, 174–198 (1996).
99. Heiney, S. A., Wohl, M. P., Chettih, S. N., Ruffolo, L. I. & Medina, J. F. Cerebellar-Dependent Expression of Motor Learning during Eyeblink Conditioning in Head-Fixed Mice. *J. Neurosci.* **34**, 14845–14853 (2014).
100. Ohyama, T. & Mauk, M. D. Cerebellar learning. in *Neurobiology of learning and memory, 2nd ed* 427–456 (Elsevier Academic Press, San Diego, CA, US, 2007). doi:10.1016/B978-012372540-0/50014-5.
101. Hansel, C., Linden, D. J. & D'Angelo, E. Beyond parallel fiber LTD: the diversity of synaptic and non-synaptic plasticity in the cerebellum. *Nat. Neurosci.* **4**, 467–475 (2001).
102. Tabata, T. & Kano, M. Synaptic Plasticity in the Cerebellum. in *Handbook of Neurochemistry and Molecular Neurobiology: Neural Signaling Mechanisms* (eds. Lajtha,

A. & Mikoshiba, K.) 63–86 (Springer US, Boston, MA, 2009). doi:10.1007/978-0-387-30370-3_6.

103. Suvrathan, A., Payne, H. L. & Raymond, J. L. Timing Rules for Synaptic Plasticity Matched to Behavioral Function. *Neuron* **92**, 959–967 (2016).

104. Payne, H. L. *et al.* Cerebellar Purkinje cells control eye movements with a rapid rate code that is invariant to spike irregularity. *eLife* **8**, e37102 (2019).

105. Lisberger, S. G. & Pavelko, T. A. Brain stem neurons in modified pathways for motor learning in the primate vestibulo-ocular reflex. *Science* **242**, 771–773 (1988).

106. Lisberger, S. G. Visual Guidance of Smooth Pursuit Eye Movements. *Annu. Rev. Vis. Sci.* **1**, 447–468 (2015).

107. Krauzlis, R. J. & Lisberger, S. G. Simple spike responses of gaze velocity Purkinje cells in the floccular lobe of the monkey during the onset and offset of pursuit eye movements. *J. Neurophysiol.* **72**, 2045–2050 (1994).

108. Crowley, J. J., Fioravante, D. & Regehr, W. G. Dynamics of Fast and Slow Inhibition from Cerebellar Golgi Cells Allow Flexible Control of Synaptic Integration. *Neuron* **63**, 843–853 (2009).

109. Kennedy, A. *et al.* A temporal basis for predicting the sensory consequences of motor commands in an electric fish. *Nat. Neurosci.* **17**, 416–422 (2014).

110. Zampini, V. *et al.* Mechanisms and functional roles of glutamatergic synapse diversity in a cerebellar circuit. *eLife* **5**, e15872 (2016).

111. Leigh, R. J. & Zee, D. S. *The Neurology of Eye Movements*. (Oxford University Press, 2015).

112. Ramachandran, R. & Lisberger, S. G. Normal Performance and Expression of Learning in the Vestibulo-Ocular Reflex (VOR) at High Frequencies. *J. Neurophysiol.* **93**, 2028–2038 (2005).

113. Robinson, D. A. A Method of Measuring Eye Movement Using a Scleral Search Coil in a Magnetic Field. *IEEE Trans. Bio-Med. Electron.* **10**, 137–145 (1963).

114. Rashbass, C. The relationship between saccadic and smooth tracking eye movements. *J. Physiol.* **159**, 326–338 (1961).

115. de Brouwer, S., Yuksel, D., Blohm, G., Missal, M. & Lefèvre, P. What Triggers Catch-Up Saccades During Visual Tracking? *J. Neurophysiol.* **87**, 1646–1650 (2002).

116. Hall, N. J., Herzfeld, D. J. & Lisberger, S. G. Evaluation and resolution of many challenges of neural spike sorting: a new sorter. *J. Neurophysiol.* **126**, 2065–2090 (2021).

117. Herzfeld, D. J. & Beardsley, S. A. Synaptic weighting for physiological responses in recurrent spiking neural networks. in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society,EMBC* 4187–4190 (2011). doi:10.1109/IEMBS.2011.6091039.

118. Ruigrok, T. J. H., Hensbroek, R. A. & Simpson, J. I. Spontaneous Activity Signatures of Morphologically Identified Interneurons in the Vestibulocerebellum. *J. Neurosci.* **31**, 712–724 (2011).

119. Walsh, J. V., Houk, J. C., Atluri, R. L. & Mugnaini, E. Synaptic Transmission at Single Glomeruli in the Turtle Cerebellum. *Science* **178**, 881–883 (1972).

120. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. Preprint at https://doi.org/10.48550/arXiv.1412.6980 (2017).

121. Loshchilov, I. & Hutter, F. SGDR: Stochastic Gradient Descent with Warm Restarts. Preprint at https://doi.org/10.48550/arXiv.1608.03983 (2017).

122. Herzfeld, D. J., Kojima, Y., Soetedjo, R. & Shadmehr, R. Encoding of error and learning to correct that error by the Purkinje cells of the cerebellum. *Nat. Neurosci.* **21**, 736–743 (2018).