

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

3 David J. Herzfeld and Stephen G. Lisberger

4

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

6

7 **Correspondence:** David J. Herzfeld, Department of Neurobiology, 311 Research Drive, Box
8 3209, Duke University School of Medicine, Durham, NC 27710, USA. Email:
9 david.herzfeld@duke.edu

10 **Running title:** Cerebellar computations during smooth pursuit

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

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

14 **Number of figures:** total, 7; color, 7

15 **Abstract**

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

29

30 **Impact statement**

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

38

39 **Introduction**

40 Understanding how neural circuits process information to generate behavior is a fundamental
41 goal of systems neuroscience. Population dynamics^{1,2} is a key feature of neural processing, and
42 the challenge of generating behavior can be envisioned as multiple transformations³ of
43 population dynamics - from sensory inputs⁴ to intermediate representations⁵ to the ultimate drive
44 of precise muscle activations⁶. A complete understanding of how neural circuits transform
45 population dynamics requires four key components: 1) precise control over the input stimuli and
46 quantifiable behavioral outputs, 2) fundamental knowledge of the necessary brain regions and
47 their underlying neuron types and circuit architecture, 3) recordings of activity from complete
48 neural populations in the requisite brain regions, and 4) the ability to link neuron types with
49 neural responses for the major neuron types within each circuit. We have achieved all four
50 components for one brain region and one sensory-motor behavior and therefore can explain how
51 a specific circuit computes, in both space and time.

52 The cerebellum offers unique advantages for understanding neural circuit computations. First,
53 the cytoarchitecture of the cerebellum is highly conserved across regions and species^{7,8}, with a
54 well-characterized circuit organization⁹ and discrete neuron types^{9,10} organized around a primary
55 three-layer feedforward network^{11,12}: incoming mossy fibers synapse onto granule cells, whose
56 parallel fiber axons provide excitatory input to Purkinje cells, the sole output of the cerebellar

57 cortex. Additional neuron types¹⁰, including Golgi cells, unipolar brush cells, and molecular
58 layer interneurons (among others^{13,14}), use both feed-forward and recurrent connections¹¹ to
59 modulate the activity within the pathway between mossy fibers and Purkinje cells. Second, many
60 cerebellar regions are crucial for the precise execution of motor behaviors^{15–19}. Motor behaviors
61 provide an ideal behavior to study neural circuit function as the parameters of movements can be
62 quantified with high spatial and temporal precision. Third, recent advances in large-scale multi-
63 contact electrode technology^{20,21} have enabled extracellular recordings from large populations of
64 well-isolated cerebellar neurons during motor behaviors^{22–27}. Finally, and crucially, our recent
65 work allows identification of many of the cerebellar neuron types directly from extracellular
66 recordings made during behavior^{23,28}. Together, the cerebellar circuit, advances in recording
67 technology, parametrizable motor behavior, and cerebellar neuron-type identification satisfy the
68 four key requirements to link neural circuit function to behavior.

69 Here, we reveal how a full cerebellar circuit computes by taking advantage of knowing the input-
70 output transformation in the cerebellar floccular complex during an exemplar motor behavior for
71 which it is essential^{18,29}, smooth pursuit eye movements. The floccular complex causally controls
72 smooth pursuit: it has disynaptic connections to extraocular motoneurons³⁰ and microstimulation
73 elicits short-latency smooth eye movements^{31,32}. Recordings from directly-identified Purkinje
74 cells^{33,34} and putative mossy fiber inputs tentatively characterize the input-output transformation
75 of the floccular circuit. Purkinje cell responses are related to eye velocity^{32,35,36} with directional
76 preferences that align with the vestibular labyrinths³⁷. Mossy fiber inputs exhibit responses
77 resembling motor corollary discharge signals, driven principally by a combination of eye
78 position and velocity^{37,38} along the horizontal and vertical axes³⁷. The requirement for seamless
79 interaction of coordinate frames dictated by vestibular sensory signals and eye movement motor
80 corollary discharge creates contrasting dynamics of floccular inputs and outputs and defines a set
81 of floccular input-output transformations.

82 Our contribution is to show how specific interneuron types work together to transform mossy
83 fiber inputs into Purkinje cell outputs through circuit function and to highlight how the floccular
84 circuit operation could generalize widely. We overcame two challenges that have prevented a
85 complete understanding of cerebellar computations from extracellular recordings to date: 1)
86 granule cells are inaccessible to conventional recording technologies²³, and 2) Purkinje cells
87 were the sole neuron type within the cerebellar circuit that could be identified definitively. Our
88 recordings from several classes of identified interneurons allow us to discriminate among
89 previously-proposed models of granule cell dynamics^{39–44}. We used computational modeling to
90 identify granule cell representations with the capacity to generate the measured responses of all
91 downstream neural populations. The result is a complete model of cerebellar circuit processing
92 that incorporates granule cell computations, accounts for downstream population dynamics, and
93 also explains several previous behavioral and neurophysiological observations during cerebellar-
94 dependent smooth pursuit learning. Thus, we believe we have identified how a complete neural
95 circuit computes to generate well-calibrated motor behavior.

96 **Results**

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

101 **The floccular complex performs active computations during pursuit**

102 We recorded extracellular action potentials from the floccular circuit (Figure 1A) while monkeys
103 performed smooth pursuit eye movements. Using expert identification²⁸ (see **Methods**), validated
104 via ground-truth optogenetic identification in mice²³, we successfully identified all major neuron
105 types schematized in Figure 1A, with the exception of granule cells. During smooth pursuit
106 tracking, monkeys fixated a central dot for a randomized interval (400-800 ms) before the target
107 started to move at a constant velocity for 650 ms. In the example trial in Figure 1B, the target
108 moved rightward at 20 deg/s. Monkeys tracked the target throughout its motion and maintained
109 eccentric fixation for an additional 350 ms to receive fluid reward. Targets moved at 10, 20, or
110 30 deg/s along the four cardinal directions.

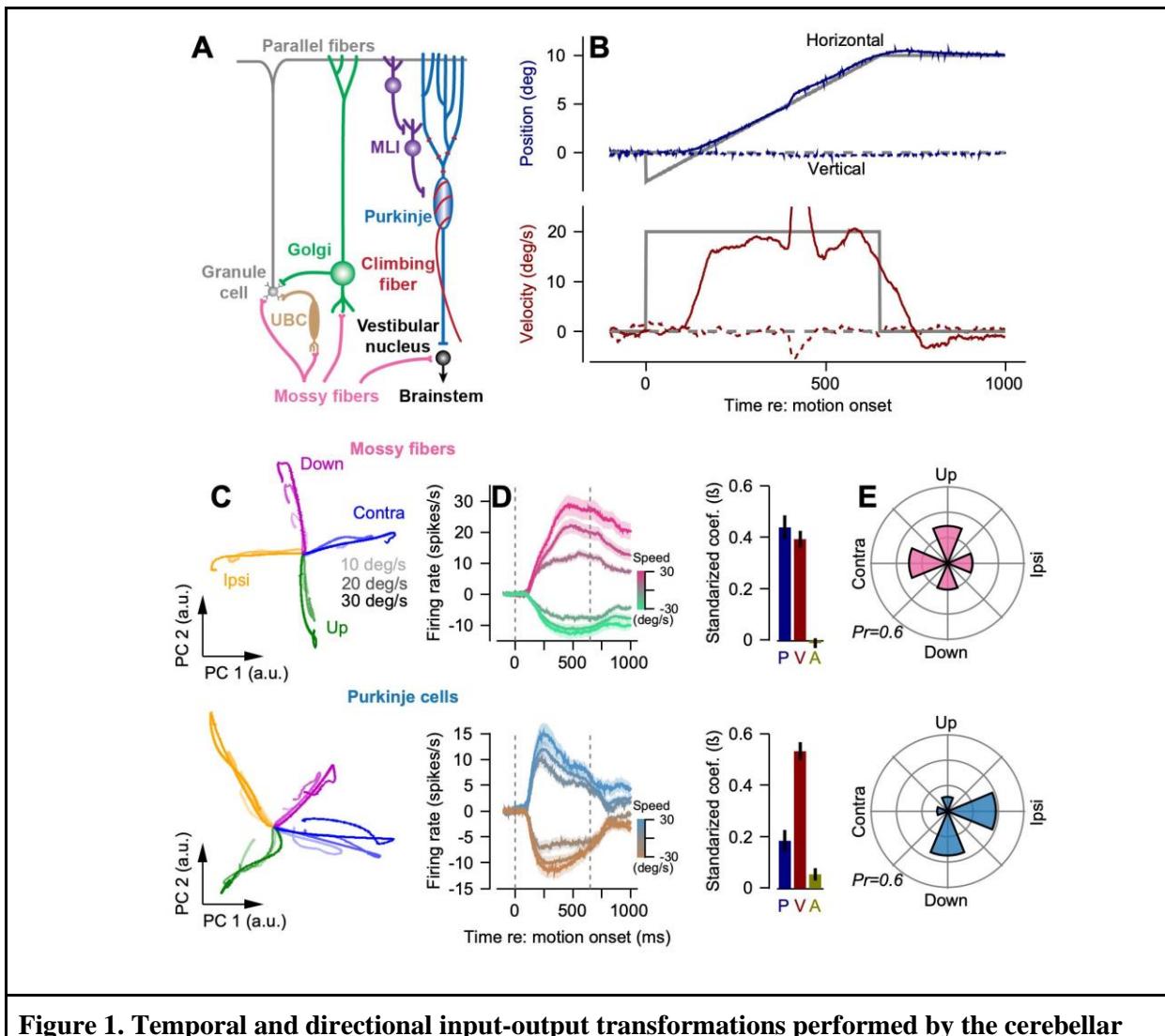


Figure 1. Temporal and directional input-output transformations performed by the cerebellar floccular complex during smooth pursuit. (A) Simplified schematic of the cerebellar circuit in the floccular complex. (B) Eye and target position (top) and velocity (bottom) versus time for an example smooth pursuit 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 \pm 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.

111
112 Separate principal component analyses of the firing rates in populations of mossy fibers and
113 Purkinje cells reveal distinct neural state-space trajectories^{1,45} (Figure 1C). Mossy fiber
114 trajectories aligned closely with the horizontal and vertical axes of the eye. The trajectories
115 terminated eccentrically in state-space with magnitudes that scaled with pursuit speed, indicating
116 persistent activity related to eccentric eye positions across all speeds and directions following
117 pursuit termination. In contrast, Purkinje cell trajectories showed principal axes that were rotated
118 relative to the cardinal axes. Their neural trajectories returned to near the origin at movement
119 completion, consistent with the absence of a relationship to eye position. For both neuron types,
120 the first two dimensions captured a substantial portion of the response variance (38% and 29%
121 for mossy fibers and 23% and 20% for Purkinje cells, a considerable amount given that we are
122 accounting for responses in 4 directions by 3 speeds of target motion).

123 To directly link neural trajectories to firing patterns, we compared mean population responses in
124 the preferred and anti-preferred directions (preferred + 180°) across all speeds (Figure 1D).
125 Mossy fiber responses resembled eye position traces, with persistent firing at movement
126 termination that scaled approximately linearly with pursuit speed (preferred direction $R^2 = 0.83 \pm$
127 0.03). The asymmetry of mossy fiber responses in the preferred versus anti-preferred directions
128 stems from complete cessation of firing in the anti-preferred direction for some mossy fibers. In
129 contrast, Purkinje cell responses peaked approximately 200 ms after motion onset with
130 magnitudes that scaled with pursuit speed (preferred direction $R^2 = 0.75 \pm 0.03$) and returned to
131 baseline at the end of pursuit regardless of final eccentric eye position. Thus, an important
132 signature of whether neurons have position-related responses is whether firing rate (1) remains
133 elevated at the end of the movement while the monkey fixates an eccentric target and (2) post-
134 movement responses scale with target speed because faster target motions make the final eye
135 position more eccentric. Regression analysis confirmed equal relations to eye position and
136 velocity with no relation from eye acceleration for mossy fibers, and a dominant relation to eye
137 velocity in Purkinje cells (histograms on right of Figure 1D). Statistical analysis verified that
138 encoding of position ($t(185)=4.08$, $p < 10^{-5}$), velocity ($t(185)=-2.86$, $p=0.005$) and acceleration
139 ($t(185)=-2.12$, $p=0.035$) all differed across mossy fibers versus Purkinje cells (independent
140 samples t-test). Documentation of the responses of each individual Purkinje cell and mossy fiber
141 appears in Supplemental Figure 1.

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

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

150 **Distinct temporal dynamics of different cerebellar neuron types**

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

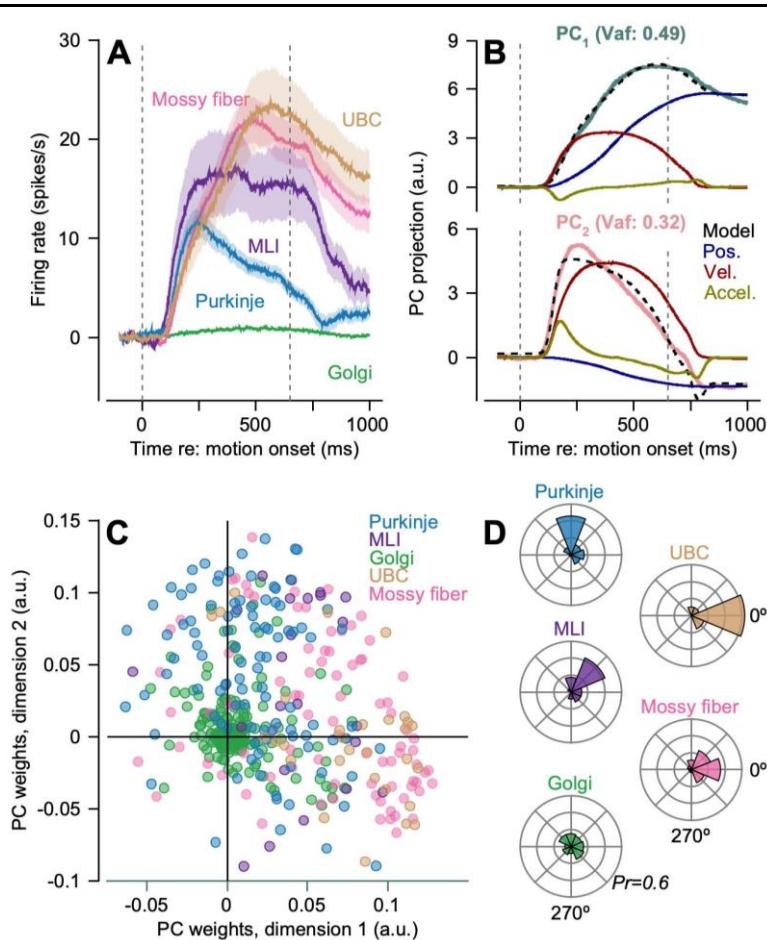


Figure 2. Functional heterogeneity of responses across cerebellar neuron types during smooth pursuit eye movements. (A) Mean modulation of firing rate as a function of time for cerebellar

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 \pm 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 (PC₁ and PC₂) 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 \pm 0.002, 0.005 \pm 0.002); Purkinje cells, (0.016 \pm 0.003, 0.042 \pm 0.005); molecular layer interneurons (0.042 \pm 0.007, 0.034 \pm 0.013); mossy fibers (0.059 \pm 0.005, 0.014 \pm 0.006); and unipolar brush cells (0.075 \pm 0.007, 0.001 \pm 0.008). (D) Distribution of angular locations, irrespective of magnitude, for each neuron type derived from the scatter plot shown in C.

161
162 Functional segregation across neuron types remained clear when analyzed at the level of
163 individual neurons instead of population means. Principal component analysis of trial-averaged
164 responses for pursuit in each neuron's preferred direction at 20 deg/s identified two dominant
165 components that together explained 81% of the population variance (49% and 32%,
166 respectively). The first component (Figure 2B, top) showed a near monotonic increase over time
167 with sustained deviation from baseline following pursuit termination. Linear regression of
168 kinematic variables showed a strong relation to eye position (standardized coefficient, $\beta = 0.84$)
169 and eye velocity ($\beta = 0.50$). The second principal component (Figure 2B, bottom) exhibited an
170 early peak followed by a return to near baseline; it was strongly related to eye velocity ($\beta = 0.83$)
171 and somewhat to eye acceleration ($\beta = 0.24$).

172 Individual neuron weights in the two-dimensional space defined by the two dominant
173 components revealed across-class distinctions along with some within-class variation (Figure
174 2C). For instance, Golgi cells are concentrated near the origin, Purkinje cells show near zero
175 weights in the first dimension but consistently positive weights in the second dimension, and
176 mossy fibers and unipolar brush cells show the opposite trend with weights centered near zero in
177 the second dimension but strongly positive in the first dimension. Permutation analysis revealed
178 that all neuron type centroids were significantly different from a neuron-type agnostic
179 distribution (mossy fibers, $n = 86$, $p = 10^{-4}$; unipolar brush cells, $n = 30$, $p = 10^{-4}$; Golgi cells, $n =$
180 186 , $p = 10^{-4}$; Purkinje cells, $n = 101$, $p = 10^{-4}$; molecular layer interneurons, $n = 23$, $p = 0.039$).
181 Much of the within-neuron-type variation in the spatial distribution shown in Figure 2C results
182 from heterogeneity in response magnitude rather than fundamental differences in temporal
183 processing within neuron types. Probability distributions of the angle of each neuron within the
184 two-dimensional space (Figure 2D) show marked differences across neuron types in the
185 distributions of the relative contributions of the first two principal components to neuron firing.
186 The distinctive neuron-type-specific response patterns highlight active computation within the
187 circuit to alter the temporal properties of the signals, with potentially distinct computational roles
188 for each neuron type.

189 **Distinct roles of cerebellar interneurons in floccular circuit processing**
190 Using our ability to identify other neuron types in the cerebellar circuit reliably^{23,28}, we next
191 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.

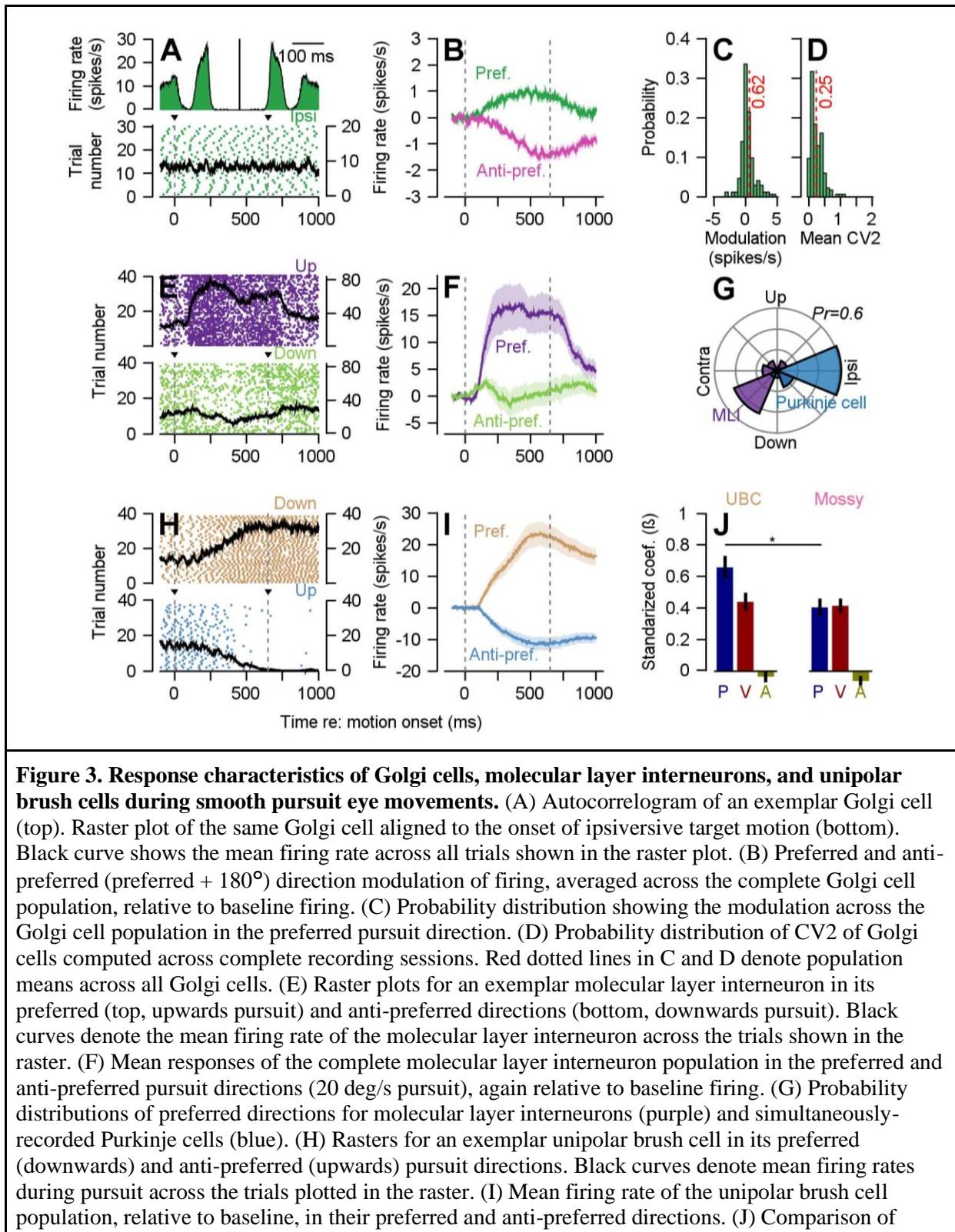


Figure 3. Response characteristics of Golgi cells, molecular layer interneurons, and unipolar brush cells during smooth pursuit eye movements. (A) Autocorrelogram of an exemplar Golgi cell (top). Raster plot of the same Golgi cell aligned to the onset of ipsiversive target motion (bottom). Black curve 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 \pm SEM across neurons.

194
195 Golgi cells are inhibitory neurons that are hypothesized to regulate granule cell activity via
196 recurrent excitation from granule cells' parallel fibers^{48–50}. Given their position in the cerebellar
197 circuit, Golgi cells could contribute to either the temporal or directional transformations. For
198 instance, Golgi cell responses consistent with eye position might subtract or block the component
199 of mossy fiber activity related to eye position via feedforward or recurrent inhibition, thereby
200 creating granule cell responses related primarily to eye velocity. Yet, we found that Golgi cells in
201 the floccular complex showed minimal modulation during smooth pursuit.

202 A typical Golgi cell showed regular firing at a low rate of ~12 spikes/s and no temporal
203 modulation of firing during ipsiversive pursuit at 20 deg/s (Figure 3A). The full Golgi cell
204 population also showed limited modulation during pursuit in their preferred or anti-preferred
205 directions at 20 deg/s (Figure 3B), quantified as the mean modulation in the preferred direction
206 (Figure 3C, 0.62 ± 0.13 spikes/s). Golgi cell firing was highly regular as measured by the mean
207 “CV2⁵¹” (Figure 3D, 0.25 ± 0.01). Only 46/186 Golgi cells showed statistically significant
208 modulation via permutation testing across pursuit directions. We conclude that Golgi cells
209 provide tonic but temporally-unmodulated inhibition of granule cells.

210 Molecular layer interneurons are positioned to perform feedforward or lateral inhibition⁵² of
211 Purkinje cells and thus could impact Purkinje cell activity both directionally and temporally. We
212 identified molecular layer interneurons by their monosynaptic inhibition of known Purkinje cells
213 in spike-timing cross-correlograms (see **Methods**). In its preferred direction, a typical molecular
214 layer interneuron's firing (Figure 3E) is related primarily to eye velocity, as it returned to near
215 baseline levels at pursuit termination, mirroring Purkinje cells. The same molecular layer
216 interneuron displayed minimal modulation of firing rate during pursuit in the anti-preferred
217 direction, a trend confirmed in averages of firing rate across the complete molecular layer
218 interneuron population (Figure 3F). The mean modulation of the molecular layer interneuron
219 population was 12.2 ± 3.0 spikes/s in their preferred direction, but only 0.4 ± 1.6 spikes/s in their
220 anti-preferred direction. The preferred direction of molecular layer interneurons was almost
221 always opposite that of their connected Purkinje cells (Figure 3G, mean angular distance: -121.8°
222 $\pm 10.4^\circ$, circular mean \pm SEM). The opposite preference for pursuit across the two populations
223 suggests that molecular layer interneurons likely contribute to the directional preferences of
224 Purkinje cells by providing inhibition in the Purkinje cell's anti-preferred direction of pursuit.

225 Unipolar brush cells are relatively common in the oculomotor regions of the cerebellar cortex,
226 but relatively rare in other cerebellar regions⁵³, suggesting a potential specialization for
227 oculomotor behavior⁵⁴. A typical unipolar brush cell (Figure 3H) showed clear positive and
228 negative modulation in its preferred and anti-preferred directions and had fully sustained
229 modulation in both directions after pursuit termination implying that they, like mossy fibers,
230 respond primarily to eye position. The same response profile appeared in the averages across the
231 complete population of unipolar brush cells, with sustained activity post-pursuit (Figure 3I).
232 Regression analysis to fit the firing rate of each neuron as a function of the simultaneously
233 recorded eye position, velocity, and acceleration showed stronger eye position encoding for

234 unipolar brush cells compared to mossy fibers (Figure 3J, standardized coefficient, $\beta = 0.65 \pm$
235 0.07 versus 0.40 ± 0.06 , $t(114) = -2.36$, $p = 0.02$). The more dominant eye position response of
236 the full population of unipolar brush cells suggests that they either perform integration of their
237 mossy fiber inputs, consistent with previous *in vitro* analyses^{55,56}, or selectively discount velocity
238 signals in their inputs from mossy fibers. Documentation of the response of each individual
239 Golgi cell, molecular layer interneuron, and unipolar brush cell appears in Supplemental Figure
240 1.

241 Golgi cells, molecular layer interneurons, and unipolar brush cells do not seem to play a role in
242 the temporal input-output transformation in the floccular complex. Golgi cells are largely
243 unmodulated, molecular layer interneurons are already related to eye velocity, and unipolar brush
244 cells are more related to position than are mossy fibers. Any role for the excitatory unipolar
245 brush cells would require subtraction from spontaneous activity that is thought not to exist in
246 granule cells⁵⁷. Finally, the temporal transformation cannot be explained by synchrony in
247 multiple mossy fibers inputs to a granule cell because we found no evidence for mossy fiber
248 synchrony other than the amount expected from covariation of mossy fiber firing rates during
249 pursuit (Supplemental Figure 2).

250 A data-driven approach for identifying granule cell transformations

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

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

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

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

276 supplied with a step input, some units produced transient, high-pass filter like responses, while
277 others showed more sustained, low-pass filter like activity (Figure 4C), reminiscent of bandpass
278 filters with varying time constants. The latency of each granule unit's peak response was
279 correlated to its response duration (correlation test between rise and decay time constants, $t(70) =$
280 6.88 , $p < 10^{-9}$). Similar patterns of units that we could characterize as high-pass and low-pass
281 were evident when we drove the model with trial-averaged mossy fiber pursuit responses (Figure
282 4D). Thus, the model network learned a range of temporal filtering characteristics where the
283 degree of filtering and subsequent duration of firing were temporally linked.

284 **A generative model of granule cell temporal transformations**

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

294 Based on the predictions of the RNN model and from iterative attempts to understand how the
295 granule cell population response would contribute to multiple features of pursuit behavior, we
296 incorporated four key features into the generative model: 1) a broad range of temporal filter
297 responses across the population so that granule units temporally tile the duration of the trial; 2)
298 strongly correlated onset and offset time constants to enable scalar variability⁶³ (i.e., “Weber
299 law”) behavior for learned timing; 3) increasing granule unit population activity with increasing
300 pursuit speeds to allow Purkinje cell firing to scale with speed; and 4) consistent timing of
301 individual granule unit activations across pursuit speed to enable generalization across pursuit
302 speeds. These characteristics are consistent with previously proposed “spectral timing” models of
303 granule cell activity^{39,64,65}.

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

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

319 population of Purkinje cells provided an excellent fit (Figure 4J, $R^2=0.92$). Thus, temporal tiling
320 in a population of granule cells can solve the position-to-velocity transformation from mossy
321 fiber to Purkinje cells. We note that there are a number of implementations of short-term
322 plasticity in the mossy fiber to granule cell synapse that could produce the temporally-tiled
323 granule cell basis set implemented by our models of granule unit activity⁶⁶. Also, other extant
324 models of granule cell responses⁴² share some of the features of the generative model and,
325 therefore, are able to reproduce the time varying Purkinje cell firing to some degree, albeit less
326 well than our model (Supplemental Figure 3).

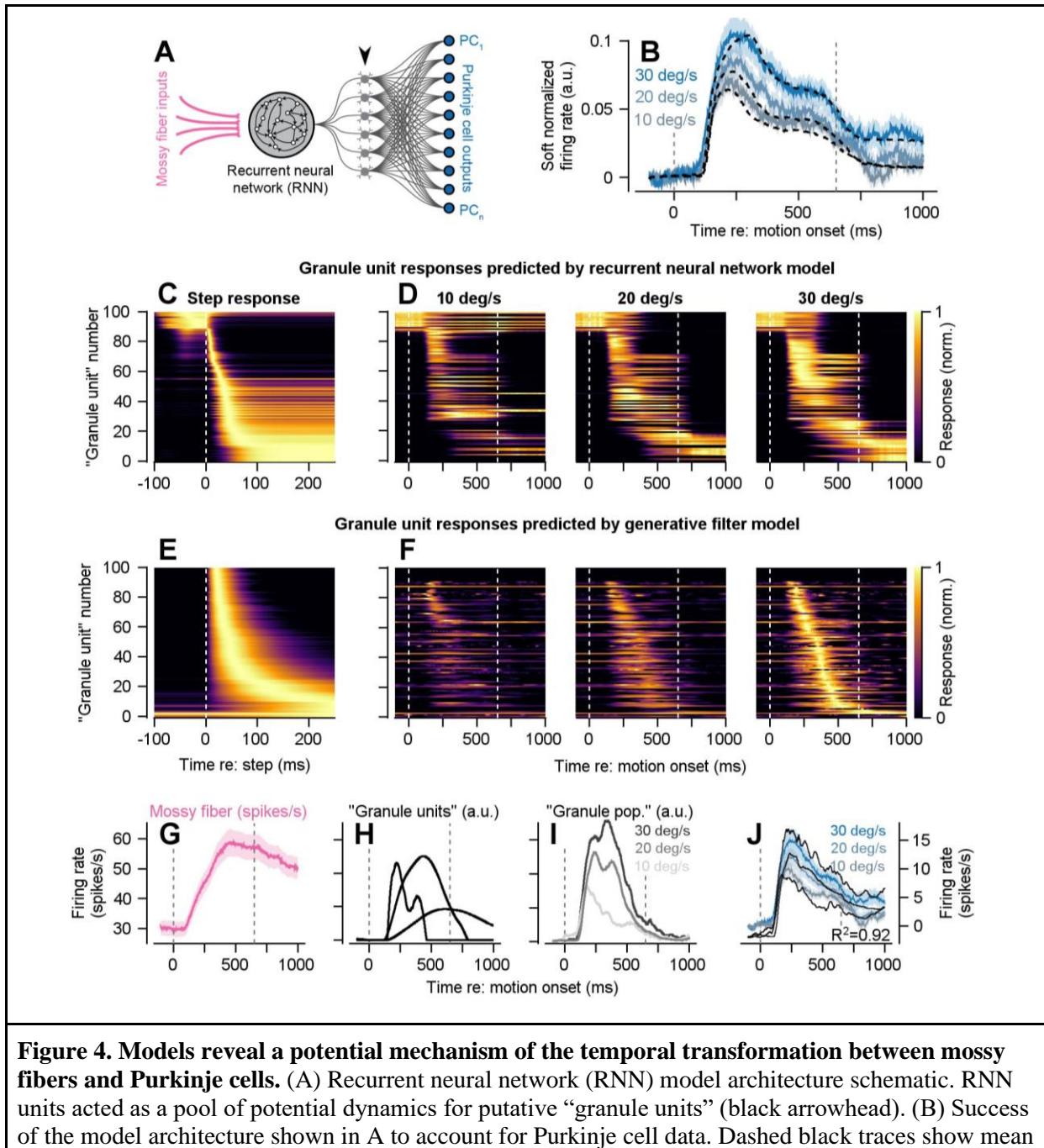


Figure 4. Models reveal a potential mechanism of the temporal transformation between mossy fibers and Purkinje cells. (A) Recurrent neural network (RNN) model architecture schematic. RNN units acted as a pool of potential dynamics for putative “granule units” (black arrowhead). (B) Success of the model 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-normalized⁴⁵ 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.

327

328

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

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

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).

347

348

349

350

351

352

353

354

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.

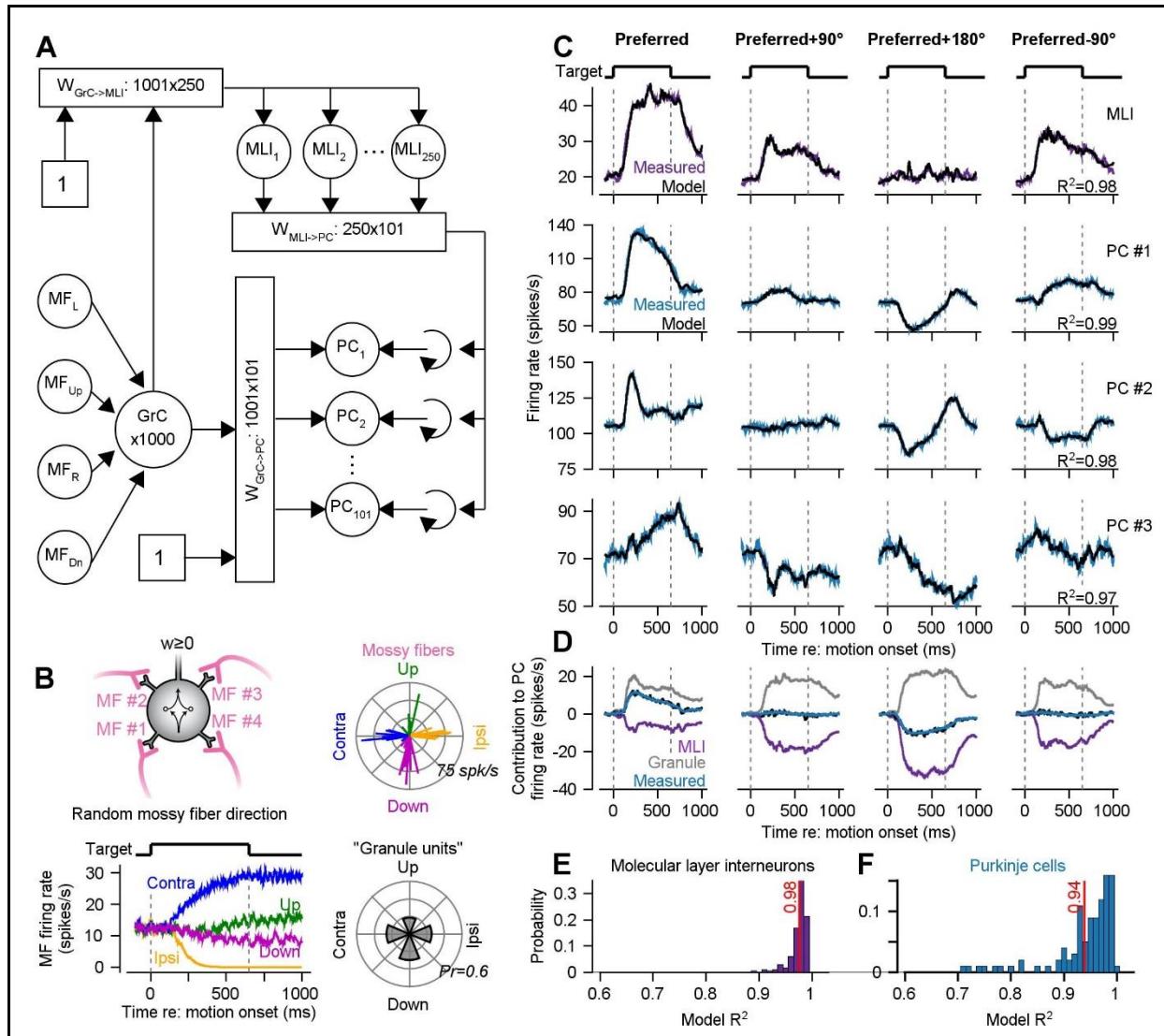


Figure 5. Model granule cell populations provide a sufficient basis for simulation of molecular layer interneuron and Purkinje cell responses. (A) Schematic diagram of a circuit model of the cerebellar cortex used to predict the responses of individual molecular layer interneurons and Purkinje cells. (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^2 's for molecular layer interneurons and (F) Purkinje cells. Vertical red lines in E-F denote the mean across the respective populations.

355

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

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

366 **Emergent properties of the cerebellar circuit model**

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

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

381 To produce directional learning in pursuit, the monkey tracks a target that moves at a constant
382 speed in an initial “pursuit direction.” After a fixed duration, an orthogonal velocity component
383 (the “instruction”) is added to the target’s motion, so that the target moves diagonally^{71,72} (Figure
384 6A, top). In the velocity domain, target motion comprises a 650-ms duration step of velocity in
385 the initial, pursuit direction (Figure 6A, middle), along with a 400-ms duration step of velocity in
386 the learning direction starting 250 ms later (Figure 6A, bottom). The instructive change in target
387 direction drives behavioral learning, assayed in the next trial or over short bouts of learning
388 trials. Learning comprises an appropriately-timed deviation of the eyes in the direction of the
389 preceding instruction that accumulates over trials (Figure 6A, arrowhead). The instruction also
390 evokes complex-spike responses in the 100 milliseconds following the instructive stimulus
391 (Figure 6D, shaded region) and causes a well-timed depression of simple-spike firing rate
392 (Figure 6G).

393 The model reproduces our previous results on generalization of single-trial learning⁷². To study
394 generalization, we induced learning with pursuit target motion at 20 deg/s and measured learning
395 in the subsequent test trial with slower or faster pursuit target motion (Figure 6B). In our
396 experiments⁷², eye movement deviation in the instruction direction of the test trial scaled with
397 test target speed in the pursuit direction (Figure 6E). We simulated the learning conditions in our
398 model by reducing the connection weights from the subset of parallel fibers active at the time of
399 instruction during pursuit at 20 deg/s (250 ms, red arrow) and measuring the learning-induced
400 changes in Purkinje cell simple-spike output on the subsequent trial for pursuit speeds of 10, 20,
401 and 30 deg/s. The model predicted a learned depression in the Purkinje cell population response

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

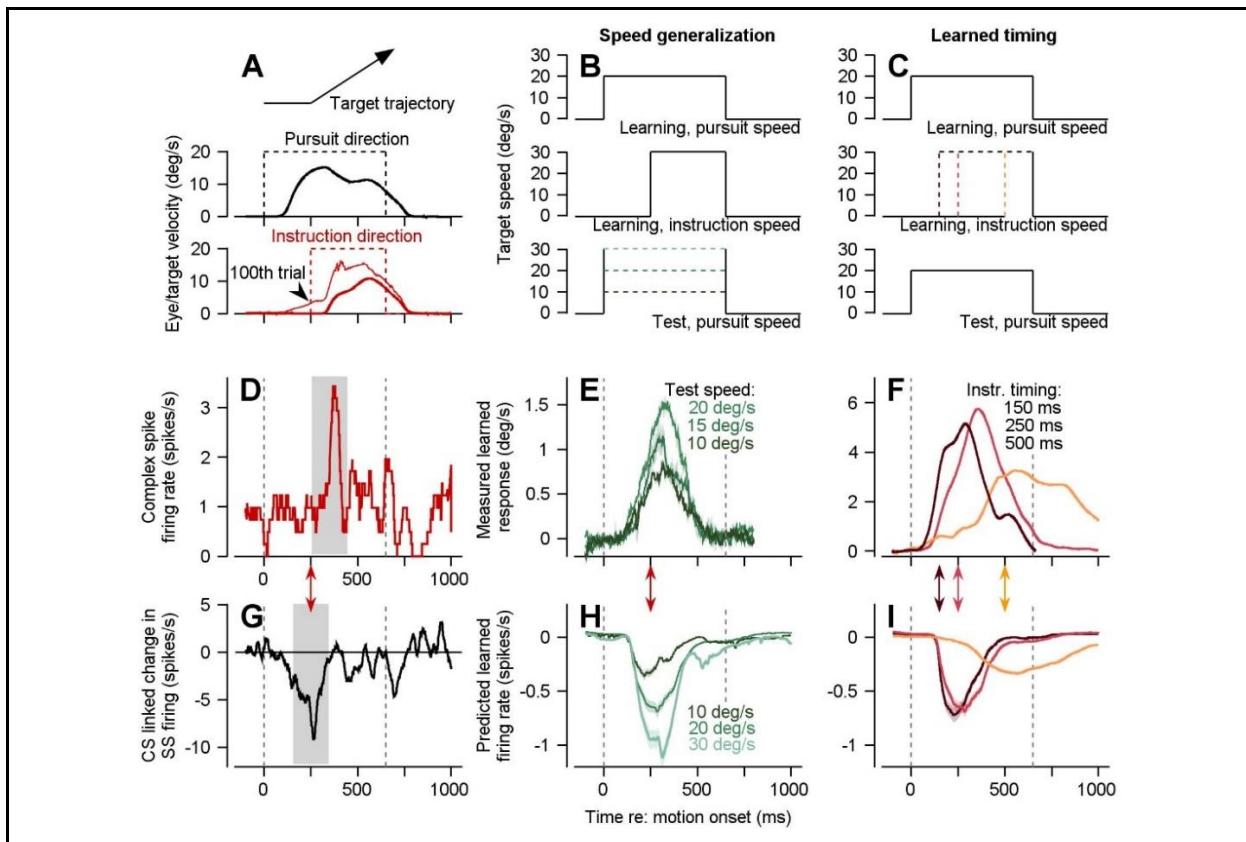


Figure 6. Emergent properties of the circuit model of the cerebellar cortex during pursuit learning. (A) Target position trajectory during a direction-learning trial (top). Eye and target velocity timeseries in the original pursuit direction (middle) and orthogonal instructive direction (bottom). Arrowhead 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 \pm SEM across neurons or behavioral learning paradigm replicates.

The model also reproduces the impact of the timing of the instructive stimulus⁷¹ 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 instruction⁷¹ (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 conditioning⁶⁶.

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 fibers⁷⁶. 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

447 firing during pursuit. In the CS-on direction, the inverse of complex-spike firing starts earlier,
448 rises faster, and decays more quickly compared to simple-spike firing (black vs. red traces,
449 bottom graph of Figure 7D). In the CS-off direction, the inverse of complex-spike firing shows
450 both an early transient increase and a positive spike after the end of target motion (cyan trace,
451 Figure 7C), neither of which has a correlate in simple-spike firing (compare black vs. cyan
452 traces, top graph Figure 7D). Thus, a linear model linking the timeseries of complex-spike
453 activity to modulation of Purkinje cell simple-spike responses across our population showed
454 limited predictive power but consistently negative correlations (Figure 7E, $r = -0.14 \pm 0.03$; $t(99)$
455 $= -5.4$, $p < 10^{-5}$ in the CS-on direction and $r = -0.24 \pm 0.02$; $t(99) = -11.3$, $p < 10^{-18}$ in the CS-off
456 direction).

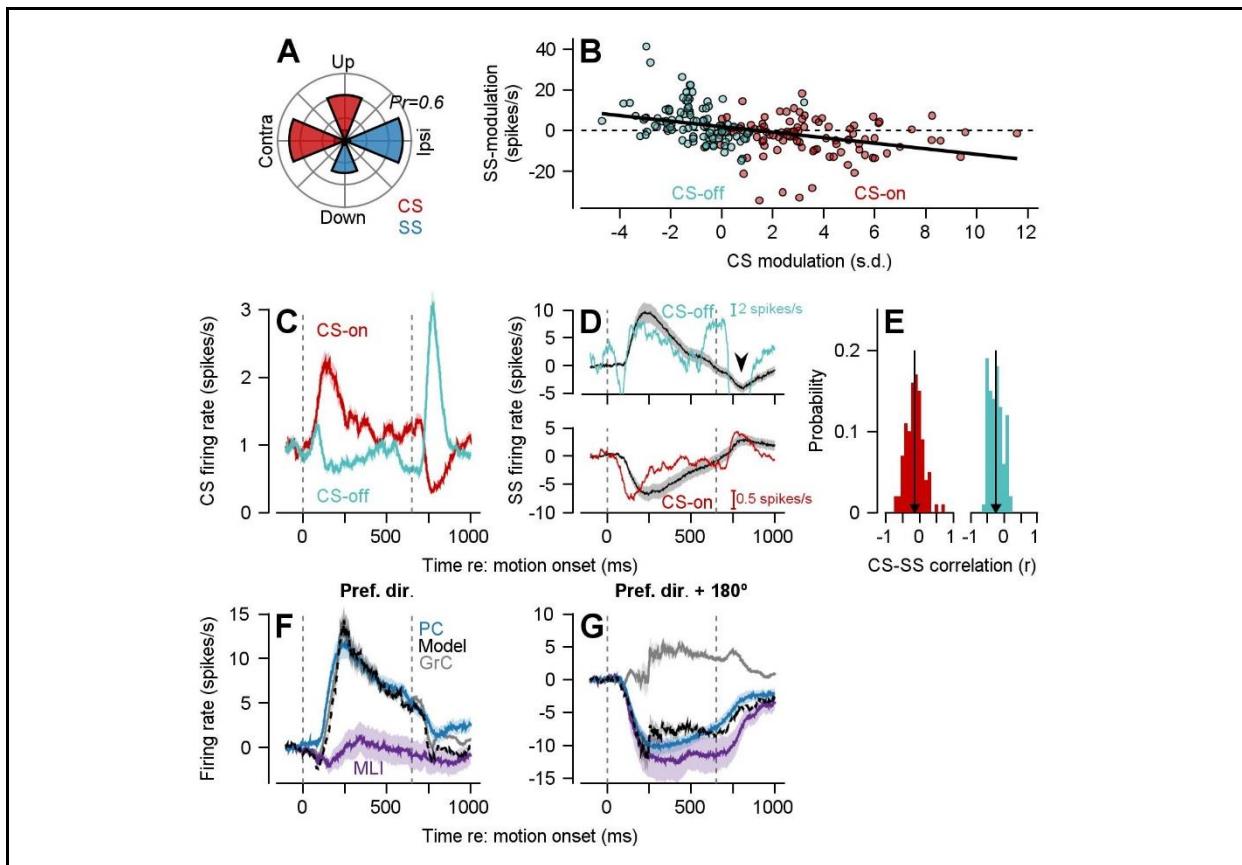


Figure 7. Evidence that complex spike mediated plasticity may create and/or maintain aspects of the temporal and directional dynamics of Purkinje cell firing. (A) Distribution of preferred pursuit directions for Purkinje cell simple spikes (blue) and their associated complex spikes (red). (B) Relationship 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 \pm SEM across neurons.

457
458 We found a tight link between the temporal properties of simple- and complex-spike firing when
459 we used the successful circuit model of Figure 5 to adapt the parallel fiber to Purkinje cell
460 synapses on the basis the complex-spike responses of our population of Purkinje cells during
461 pursuit. To obtain the fits in Figure 7F and G, we decreased the weights of individual granule
462 cell inputs to Purkinje cells based on the product of each granule unit's response during pursuit
463 and the mean complex-spike response measured in the same direction. From the excitatory input
464 of granule cells, we subtracted the weighted input of molecular layer interneurons. Thus, the fits
465 shown in Figure 7F and G had two unknowns: the magnitude of parallel fiber synaptic
466 depression due to a complex spike and the weight of inputs from molecular layer interneurons.
467
468 The excellent match between observed and predicted mean simple-spike firing (Figure 7F and G)
469 implies that climbing-fiber inputs coupled with granule layer activity could contribute to the
470 temporal dynamics of cerebellar output. As before, the success of the predictions depends on 1)
471 the temporal basis set in the granule unit population, 2) refinement of the strength of inputs from
472 the granule layer basis set by complex-spike mediated plasticity, and 3) direction-tuned
473 inhibition from molecular layer interneurons to create decreases in simple-spike firing because
474 plasticity of the inputs from spontaneously-silent granule cell inputs cannot drive Purkinje cell
responses below baseline.

475
476 **Discussion**
477 The brain generates behavior through the operation of neural circuits that transform the dynamics
478 of their inputs^{79–81} to generate outputs that facilitate appropriate downstream neural processing⁸².
479 As neural circuits are comprised of multiple neuron types^{9,83–90}, a full understanding of the nature
480 of the transformations mandates a multi-level understanding of the “parts list” of a brain area via
481 both identification of the characteristic neuron types and description of their functional
482 characteristics during behavior. Here, we demonstrate that discrete neuron types within a well-
483 defined neural circuit have distinct functional roles that allow the circuit to perform fundamental
484 transformations of population dynamics to facilitate behavioral output. We also show how
485 complex-spike mediated plasticity complements the distinct neuron-type computations to
486 facilitate the transformations between mossy fiber inputs and eventual Purkinje cell outputs.

487 **Computation by a neural circuit**
488 To understand how a specific circuit computes, we took advantage of new technology that
489 allowed us to identify neuron types in the cerebellar cortex from extracellular recordings^{23,28}. The
490 canonical cerebellar circuit, including its intrinsic interneurons, has a conserved architecture
491 across regions and species^{7,8} and supports a wide range of behaviors ranging from motor
492 control¹⁵ to communication^{91,92} to cognition^{93,94}. We recorded the temporal dynamics and
493 directional organization of firing in identified mossy fiber input elements and Purkinje cell

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

498 We suggest that an input-output transformation of temporal dynamics occurs in the cerebellar
499 input layer, potentially via synaptic⁶⁶ or granule cell intrinsic^{95,96} mechanisms. A directional
500 transformation occurs through circuit mechanisms in the last-order inputs to the output neurons.
501 We confirmed the veracity of the proposed computation through simulation of a circuit model
502 that predicts the responses of each individual Purkinje cell in our sample. Further, the model has
503 multiple emergent properties that agree with known experimental observations^{71,72}. We think that
504 we have explained how the cerebellar circuit computes during a specific sensory-motor behavior
505 that requires the region of the cerebellum we studied, namely the floccular complex.

506 The properties of the temporal basis set provided by the granule unit population were critical for
507 the success of our model. They were: 1) a broad range of temporal filter responses such that
508 granule units temporally tile complete trial; 2) strongly correlated onset and offset time constants
509 resulting in scalar variability⁶³; 3) granule unit population activity that increased with increasing
510 pursuit speeds; and 4) consistent timing of individual granule unit activations across pursuit
511 speed to enable generalization of learning across pursuit speeds. These characteristics are
512 consistent with previously proposed “spectral timing” models^{39,64,65}, and allow us to effectively
513 discount alternative models of granule unit processing⁶⁴. We can make quite specific predictions
514 about granule cell activity because (1) we measure the mossy fiber inputs and (2) we have far
515 more degrees of behavioral freedom than other tasks used to place limits on granule layer
516 processing to-date. For example, cerebellar-dependent eyelid conditioning has a single primary
517 degree of freedom: the timing between the conditioned and unconditioned stimuli. We can vary
518 timing, pursuit speed, and pursuit direction, enabling more strenuous constraints on our model of
519 circuit processing.

520 Our analysis provides a roadmap for asking whether the cerebellar cortex performs a “universal
521 transform” that aligns with the universal architecture of the cerebellar circuit^{97,98}. For example,
522 there is enough data about the responses of neurons in the eye blink area in mice⁹⁹ or rabbits¹⁰⁰ to
523 deploy our analysis approach in the eye blink cerebellar region and ask whether a universal
524 transform exists across these two behaviors. More generally, it seems plausible that many or all
525 regions of the cerebellar cortex compute temporal transformations, and the temporal basis set
526 provided by our model granule units could support many different temporal transformations on
527 the time scale of movements. Further, the multiple plasticity mechanisms^{101,102} in the cerebellar
528 cortex could adapt the transformations so that they are specialized across behaviors and areas.
529 Meta-plasticity¹⁰³ combined with the general nature of the granule unit basis set may provide an
530 even broader ability for the cerebellum to provide region-specific and task-specific circuit
531 transformations.

532 Our strategy for understanding circuit transformations certainly generalizes to different
533 cerebellar regions and likely to other brain areas. Its key features were: 1) an ability to map
534 extracellular signals to cerebellar neuron types²⁸ that is validated across species for multiple
535 regions of the cerebellum²³, 2) interpretational context that provides an understanding of how
536 modulation of firing rate in floccular Purkinje cells alters eye movements^{104,105}, 3) a region that

537 performs quantifiable temporal and directional transformations of its inputs within a broader
538 brain network that is understood quite thoroughly¹⁰⁶, 4) quantitative behavioral paradigms to
539 characterize the temporal and spatial organization of inputs, outputs, and internal processing, and
540 5) computational analyses to validate the proposed circuit model.

541 **Specifics of floccular circuit computation**

542 **Temporal processing.** In the floccular complex, mossy fiber inputs signal eye position with a
543 modest relation to eye velocity³⁸; Purkinje cell outputs are related to eye velocity¹⁰⁷ with a
544 modest relation to eye acceleration³⁵. Our data imply that the input-output transformation of
545 temporal dynamics occurs in the granule cell layer, potentially via cellular mechanisms at the
546 mossy fiber to granule cell synapse⁶⁶. We localized the transformation to that synapse through
547 the observation that all downstream neurons, including Golgi cells, molecular layer interneurons,
548 and Purkinje cells, show evidence of the full temporal transformation.

549 Our analyses suggest the role of two kinds of granule layer interneurons, Golgi cells and unipolar
550 brush cells, in distinct aspects of granule cell activity. While we predicted *a priori* that Golgi
551 cells would be instrumental in actively performing temporal transformations, and they may be in
552 other systems, they proved to be only weakly temporally modulated (in relation to eye velocity)
553 in our system. They seem to provide tonic inhibition that may help to create a sparse but
554 temporally diverse¹⁰⁸ representation in the granule cells. Consistent with *in vitro* results^{55,56},
555 unipolar brush cells seem to temporally integrate the characteristics of upstream mossy fiber
556 inputs, potentially diversifying the temporal inputs to granule cells^{109,110}.

557 An engineer would not perform temporal transformations by creating a temporally-sparse basis
558 set as an intermediate population code that transforms sustained inputs into a temporally-
559 distributed set of transient responses. Yet, biology may use that strategy because it allows a
560 broad range of transformations of temporal dynamics, not just mathematical differentiation, and
561 has the added advantage of affording temporally-specific learning of sensory-motor
562 transformations. Sadly, current extracellular recording technology does not yet allow access to
563 granule cell responses²³. Thus, our theory, like others, makes a prediction that will need to be
564 tested once technology has evolved.

565 **Directional processing.** In the floccular complex, mossy fiber inputs are organized according to
566 the cardinal directions – left, right, up, down – while Purkinje cell outputs prefer “ipsiversive”
567 (towards the side of recording) or downward plus slightly contraversive eye motion³⁶. Our
568 analysis of the direction tuning in different neuron types implies that parallel fibers with all
569 direction tunings cross the dendrites of individual Purkinje cells. Multi-directional signals are
570 necessary in the parallel fiber inputs to each Purkinje cell, both to allow cerebellar learning to
571 tune the directionality of Purkinje cells and to enable directional learning in pursuit⁷². Thus,
572 unlike the temporal transformation in the input layer, the directional transformation in the
573 floccular complex appears to occur in the molecular layer.

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

579 firing^{46,47,73,75} through complex-spike mediated synaptic depression of directional parallel fiber to
580 Purkinje cell synapses. We note that our statements about molecular layer interneurons are
581 limited to the class called “MLI-1”, which inhibit Purkinje cells²⁷ and constitute our entire
582 expert-identified sample. We cannot make any statements about the class called “MLI-2”, which
583 inhibit other MLIs but were not included in our sample.

584 ***Functional significance of the input-output transformations.*** Why is it necessary to
585 functionally differentiate the signals of mossy fiber inputs to the floccular complex? We assume
586 that the inputs to the flocculus are limited by the “language” spoken by the brainstem¹⁰⁷, where a
587 “neural integrator” transforms transient inputs into the sustained firing needed to maintain stable
588 eye position at any eccentric position¹¹¹. Thus, the corollary discharge inputs to the floccular
589 complex signal eye position. Yet, the output from the floccular complex must interface in the
590 brainstem with the head-velocity-related inputs from the vestibular system and thus must signal
591 eye velocity. The cerebellum’s transformations need to adapt to the fixed and immutable
592 demands of the sensory and motor periphery.

593 **Limitations of the study**

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

605

606 **Acknowledgements**

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

611 **Author contributions**

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

615 **Conflicts of interest**

616 The authors declare no conflicts of interest.

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

624 **General procedures**
625 Prior to behavioral training or neurophysiological recordings, monkeys were instrumented with a
626 head-restraint device, an eye coil, and at least one recording cylinder via separate surgical
627 procedures, all performed with aseptic technique. For each surgical procedure, animals were
628 deeply anaesthetized with isoflurane. Animals received peri- and post-operative analgesia until
629 they had completely recovered. In the first of multiple surgical procedures, head restraint
630 hardware was attached to the animal's skull to allow measurement of the animal's eye position
631 without concomitant head movements. In a later surgery, we sutured a small coil of wire to the
632 sclera of one eye¹¹², allowing measurement of the animal's eye position using the search coil
633 technique¹¹³. Following these two surgical procedures, animals were trained to perform smooth
634 pursuit eye movements during discrete trials of target motion in exchange for fluid reward. The
635 experimental paradigm is described in detail below. Once the animal had demonstrated excellent
636 smooth pursuit tracking abilities, as evidenced by minimal intervening saccadic eye movements,
637 we affixed a stainless-steel recording cylinder above a craniectomy, allowing access to the
638 floccular complex of the cerebellum with microelectrodes. The position of this cylinder was 11
639 mm lateral to the midline, angled at 26° relative to the frontal plane, and pointed towards the
640 interaural axis.

641 **Behavioral task**
642 Each day animals were seated in a dark room, 30 cm in front of a CRT monitor (80 Hz refresh
643 rate, 2304 x 1440 pixels, 480 x 310 mm). Their heads were attached to the chair via the
644 previously implanted restraint hardware. Horizontal and vertical eye position signals were
645 separately digitized at 1,000 Hz using the search coil system. We computed the velocity of the
646 animal's eye movement offline using a 2nd order causal Butterworth low-pass filter with a cutoff
647 frequency of 10 Hz. As we were principally interested in the relationship between floccular
648 neural responses and smooth pursuit eye movements, we removed any saccadic eye movements
649 from ongoing pursuit using an automated procedure²². The occurrence of a saccade was
650 identified offline using eye velocity (50 deg/s) and eye acceleration thresholds (1,250 deg/s²).
651 Onset and offset of each saccade were determined as the time when the animal's eye velocity or
652 acceleration fell below both thresholds for more than 10 ms. In all analyses, we treated eye
653 kinematics during saccades as missing data.

654 Stimulus presentation was controlled via our lab's custom "Maestro" software. The visual
655 stimulus in all trials was a small (0.5° diameter) black spot shown on a light gray background. In
656 a small subset of trials, animals were required to fixate a stationary target placed at one of nine
657 locations, evenly spaced within a 10° x 10° square. The monkey was required to fixate the dot at
658 each position (within ±1°) continuously for one second in exchange for reward. In the majority
659 of trials, animals tracked smoothly moving targets. Each smooth pursuit trial began by fixating
660 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 paradigm¹¹⁴. 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 system¹¹⁵. 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 movements¹⁸. 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 (4th 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 Pursuit¹¹⁶ (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 $0.6 \pm 2.5\%$ (mean \pm 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

704 to the level of background noise measured on the same channel (the primary contact). The mean
705 signal-to-noise ratio, defined as the amplitude of the spike divided by an estimate of noise
706 computed as 1.96-times the standard deviation of background noise, was 5.6 ± 2.9 (mean \pm SD).
707 We converted neuron spike trains into firing rates by causally convolving each spike train with a
708 dual exponential kernel. The kernel ($\tau_{\text{rise}} = 0.1$ ms and $\tau_{\text{decay}} = 50$ ms) mimics the properties of
709 post-synaptic currents¹¹⁷ while preserving the ability to measure latencies with high precision.
710 Autocorrelograms were computed using previously described techniques²², and normalized by
711 the bin width (1 ms), placing autocorrelograms in units of spikes/s.

712 **Expert identification of cerebellar neuron types**

713 A detailed discussion of the methodology for identification of cerebellar neuron types in our
714 recordings has been published previously²⁸. Briefly, we began by identifying a subset of units as
715 ground-truth Purkinje cells based on their unique extracellular properties. Purkinje cells receive
716 excitatory inputs from parallel fibers as well as strong inputs via climbing fibers from the inferior
717 olive that drive post-synaptic Purkinje cell complex spikes. During complex spikes and for
718 several milliseconds thereafter, Purkinje cells are prevented from firing simple spikes, resulting
719 in a stereotypical complex-spike-induced pause in simple spikes. We included only Purkinje
720 cells that had a complex-spike-induced pause.

721 We identified molecular layer interneurons by establishing monosynaptic functional inhibitory
722 relationships between these neurons and simultaneously recorded ground-truth Purkinje cells.
723 Therefore, our sample of molecular layer interneurons constitutes principally MLI-1's, classified
724 by others based on their genetic¹³ and connection profiles²⁷. Golgi cells were identified by their
725 low firing rate responses, characteristic extracellular waveforms¹¹⁸, and presence within the
726 granule cell layer. Mossy fiber inputs to the cerebellar cortex were identified based on the
727 presence of a negative-after-wave (NAW). The NAW corresponds to the postsynaptic response
728 of granule cells within the cerebellar glomerulus¹¹⁹. We note that while a NAW is sufficient to
729 identify mossy fibers, it is possible to record from mossy fiber axons that do not show marked
730 NAWs. To ensure that our sample contained only known mossy fibers, we excluded from
731 analysis putative mossy fibers that lacked a NAW. Unipolar brush cells were identified based on
732 established functional responses recorded by *in vitro* studies showing that unipolar brush cells
733 elongate the timescales of discrete mossy fiber input over 10s to hundreds of milliseconds^{55,56}.
734 We considered units to be unipolar brush cells if triggering their firing off simultaneously
735 recorded mossy fiber bursts yielded temporally-elongated responses. We recently validated our
736 expert-labeling approach using a deep-learning classifier trained on ground-truth
737 optogenetically-identified recordings in mice²³.

738 **Principal component analysis**

739 To identify the primary modes of temporal information contained in the cerebellar population
740 during smooth pursuit eye movements, we performed principal component analysis across all
741 expert-identified neurons. As the responsiveness of cerebellar neurons to pursuit varied widely,
742 we wanted to ensure that highly responsive neurons did not bias our estimate of the principal
743 component directions. Therefore, we performed soft-normalization of neuron firing rates⁴⁵,
744 measured in their preferred direction, prior to principal component analysis. Soft-normalization
745 squashes modulations less than 5 spikes/s to timeseries near zero and scales modulations of
746 greater than 5 spikes/s to approximately unit magnitude. After soft-normalization, we smoothed
747 the firing rate traces using a boxcar filter with a 50 ms width. To identify the principal

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

756 **Regression of neuron firing rates to eye kinematics**

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

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

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

772 **Recurrent neural network model of granule cell transformations**

773 We constructed a recurrent neural network model to investigate features of the transformation
774 between mossy fiber inputs and Purkinje cell outputs. We supplied as inputs to the model single-
775 trial firing rate responses of individual mossy fibers ($n = 86$) during 10, 20, or 30 deg/s smooth
776 pursuit trials in the preferred direction of each mossy fiber. We trained the complete network
777 model to predict the single-trial firing responses of our recorded Purkinje cell population.
778 Purkinje cell responses were in their preferred directions with speeds that were trial-matched to
779 the speed of pursuit supplied as mossy fiber inputs (e.g., mossy fiber inputs corresponding to
780 single-trial responses during 10 deg/s target motion were paired with Purkinje cell outputs also
781 measured during 10 deg/s pursuit). To ease training, we preconditioned the input and output data
782 by soft-normalizing⁴⁵ the responses of the mossy fiber and Purkinje cell responses. The soft-
783 normalized mossy fiber inputs were supplied to a pool of RNN units, whose dynamics were
784 governed by:

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

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

787 (dimensions 25 x 25) serves as recurrent weights on RNN unit activity from the previous
788 timestep while the matrix W_i (dimensions 25 x 86) weighs incoming mossy fiber inputs,
789 represented by $u(t)$. The vector b serves as a time-invariant bias. The combination of inputs from
790 upstream mossy fiber activity, recurrent inputs, and the bias parameter was passed through a
791 non-linear activation function, σ , chosen to be the sigmoid function. The reservoir of RNN units
792 served as the temporal bases for the subsequent “granule unit” layer. Each granule unit received
793 a weighted set of inputs from the 25 RNN units. Granule unit inputs were passed through a
794 sigmoid activation function, constraining the activity of individual granule units between zero
795 and one. Finally, a set of fully connected weights mapped the $n = 100$ granule units onto each
796 output Purkinje cell. Unlike previous layers, we constrained the weight matrix between granule
797 units and Purkinje cells to be strictly positive. The complete network was simultaneously trained
798 using the Adam¹²⁰ optimizer with cosine annealing with warm-restarts¹²¹. We used drop-out
799 layers during training with 50% probability of dropout before and after the granule unit layer to
800 avoid overfitting and ensure that the dynamics necessary to represent Purkinje cell firing rates
801 were distributed throughout the granule layer units. We used early termination to stop training
802 when the cross-validated error from a withheld 10% of the single trial data failed to decrease for
803 more than 10 training epochs.

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

814 A generative model of granule cell dynamics

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

825 We began by establishing synaptic input to each granule unit:

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

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

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

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

850 **Cerebellar circuit model for predicting downstream neural responses**

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

$$J_n = (\mathbf{y}_n^{MLI} - R^T \mathbf{w}_n)^T (\mathbf{y}_n^{MLI} - R^T \mathbf{w}_n) + \lambda \mathbf{w}_n^T \text{diag} \left[\frac{1}{\sigma^2(R)} \right] \mathbf{w}_n \quad (5)$$

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

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

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

885 **Evaluation of emergent properties of the cerebellar circuit model**

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

$$\Delta\mathbf{w}_n = -k \frac{\mathbf{r}_n(\text{Instr.}+75\text{ms})}{\max [\mathbf{r}_n(\text{Instr.}+75\text{ms})]} \quad (6)$$

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

902 **Statistical analysis**

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

908 **Data availability**

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

911 **References**

- 912 1. Vyas, S., Golub, M. D., Sussillo, D. & Shenoy, K. V. Computation Through Neural
913 Population Dynamics. *Annu. Rev. Neurosci.* **43**, 249–275 (2020).
- 914 2. Shenoy, K. V. & Kao, J. C. Measurement, manipulation and modeling of brain-wide neural
915 population dynamics. *Nat. Commun.* **12**, 633 (2021).
- 916 3. Khilkevich, A. *et al.* Brain-wide dynamics linking sensation to action during decision-
917 making. *Nature* **634**, 890–900 (2024).
- 918 4. Runyan, C. A., Piasini, E., Panzeri, S. & Harvey, C. D. Distinct timescales of population
919 coding across cortex. *Nature* **548**, 92–96 (2017).
- 920 5. Gold, J. I. & Shadlen, M. N. The Neural Basis of Decision Making. *Annu. Rev. Neurosci.* **30**,
921 535–574 (2007).
- 922 6. Inagaki, H. K. *et al.* Neural Algorithms and Circuits for Motor Planning. *Annu. Rev.
923 Neurosci.* **45**, 249–271 (2022).
- 924 7. Striedter, G. F., Bullock, T. H., Preuss, T. M., Rubenstein, J. & Krubitzer, L. A. *Evolution of
925 Nervous Systems*. (Academic Press, 2016).
- 926 8. Jacobs, B. *et al.* Comparative neuronal morphology of the cerebellar cortex in afrotherians,
927 carnivores, cetartiodactyls, and primates. *Front. Neuroanat.* **8**, (2014).
- 928 9. Ramón y Cajal, S. *Histologie Du Système Nerveux de l'homme & Des Vertébrés*. (Maloine,
929 Paris, 1909). doi:10.5962/bhl.title.48637.
- 930 10. Hull, C. & Regehr, W. G. The Cerebellar Cortex. *Annu. Rev. Neurosci.* **45**, 151–175 (2022).
- 931 11. De Zeeuw, C. I., Lisberger, S. G. & Raymond, J. L. Diversity and dynamism in the
932 cerebellum. *Nat. Neurosci.* **24**, 160–167 (2021).
- 933 12. Eccles, J. C., Ito, M. & Szentágothai, J. *The Cerebellum as a Neuronal Machine*. (Springer,
934 Berlin, Heidelberg, 1967). doi:10.1007/978-3-662-13147-3.
- 935 13. Kozareva, V. *et al.* A transcriptomic atlas of mouse cerebellar cortex comprehensively
936 defines cell types. *Nature* **598**, 214–219 (2021).
- 937 14. Osorno, T. *et al.* Candelabrum cells are ubiquitous cerebellar cortex interneurons with
938 specialized circuit properties. *Nat. Neurosci.* **25**, 702–713 (2022).
- 939 15. Holmes, G. The Symptoms of Acute Cerebellar Injuries due to Gunshot Injuries. *Brain* **40**,
940 461–535 (1917).
- 941 16. Barash, S. *et al.* Saccadic Dysmetria and Adaptation after Lesions of the Cerebellar Cortex.
942 *J. Neurosci.* **19**, 10931–10939 (1999).
- 943 17. Ritchie, L. Effects of cerebellar lesions on saccadic eye movements. *J. Neurophysiol.* **39**,
944 1246–1256 (1976).
- 945 18. Rambold, H., Churchland, A., Selig, Y., Jasmin, L. & Lisberger, S. G. Partial Ablations of
946 the Flocculus and Ventral Paraflocculus in Monkeys Cause Linked Deficits in Smooth
947 Pursuit Eye Movements and Adaptive Modification of the VOR. *J. Neurophysiol.* **87**, 912–
948 924 (2002).
- 949 19. Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J. & Thach, W. T. Throwing while
950 looking through prisms. *Brain* **119**, 1183–1198 (1996).
- 951 20. Steinmetz, N. A. *et al.* Neuropixels 2.0: A miniaturized high-density probe for stable, long-
952 term brain recordings. *Science* **372**, eabf4588 (2021).
- 953 21. Jun, J. J. *et al.* Fully integrated silicon probes for high-density recording of neural activity.
954 *Nature* **551**, 232–236 (2017).
- 955 22. Herzfeld, D. J., Joshua, M. & Lisberger, S. G. Rate versus synchrony codes for cerebellar
956 control of motor behavior. *Neuron* **111**, 2448–2460.e6 (2023).

- 957 23. Beau, M. *et al.* A deep-learning strategy to identify cell types across species from high-
958 density extracellular recordings. 2024.01.30.577845 Preprint at
959 <https://doi.org/10.1101/2024.01.30.577845> (2024).
- 960 24. Kostadinov, D., Beau, M., Blanco-Pozo, M. & Häusser, M. Predictive and reactive reward
961 signals conveyed by climbing fiber inputs to cerebellar Purkinje cells. *Nat. Neurosci.* **22**,
962 950–962 (2019).
- 963 25. Sedaghat-Nejad, E., Pi, J. S., Hage, P., Fakharian, M. A. & Shadmehr, R. Synchronous
964 spiking of cerebellar Purkinje cells during control of movements. *Proc. Natl. Acad. Sci.* **119**,
965 e2118954119 (2022).
- 966 26. Pi, J. S. *et al.* The olfactory input to the cerebellum dissociates sensory events from movement
967 plans. *Proc. Natl. Acad. Sci.* **121**, e2318849121 (2024).
- 968 27. Lackey, E. P. *et al.* Specialized connectivity of molecular layer interneuron subtypes leads to
969 disinhibition and synchronous inhibition of cerebellar Purkinje cells. *Neuron* **0**, (2024).
- 970 28. Herzfeld, D. J., Hall, N. J. & Lisberger, S. G. Strategies to decipher neuron identity from
971 extracellular recordings in the cerebellum of behaving non-human primates.
972 2025.01.29.634860 Preprint at <https://doi.org/10.1101/2025.01.29.634860> (2025).
- 973 29. Belton, T. & McCrea, R. A. Role of the cerebellar flocculus region in the coordination of eye
974 and head movements during gaze pursuit. *J. Neurophysiol.* **84**, 1614–1626 (2000).
- 975 30. Highstein, S. M. The organization of the vestibulo-oculomotor and trochlear reflex pathways
976 in the rabbit. *Exp. Brain Res.* **17**, 285–300 (1973).
- 977 31. Lisberger, S. G., Pavelko, T. A. & Broussard, D. M. Responses during eye movements of
978 brain stem neurons that receive monosynaptic inhibition from the flocculus and ventral
979 paraflocculus in monkeys. *J. Neurophysiol.* **72**, 909–927 (1994).
- 980 32. Joshua, M., Medina, J. F. & Lisberger, S. G. Diversity of Neural Responses in the Brainstem
981 during Smooth Pursuit Eye Movements Constrains the Circuit Mechanisms of Neural
982 Integration. *J. Neurosci.* **33**, 6633–6647 (2013).
- 983 33. Bell, C. C. & Grimm, R. J. Discharge properties of Purkinje cells recorded on single and
984 double microelectrodes. *J. Neurophysiol.* **32**, 1044–1055 (1969).
- 985 34. Bloedel, J. R. & Roberts, W. J. Action of climbing fibers in cerebellar cortex of the cat. *J.*
986 *Neurophysiol.* **34**, 17–31 (1971).
- 987 35. Lisberger, S. G. & Fuchs, A. F. Role of primate flocculus during rapid behavioral
988 modification of vestibuloocular reflex. I. Purkinje cell activity during visually guided
989 horizontal smooth-pursuit eye movements and passive head rotation. *J. Neurophysiol.* **41**,
990 733–763 (1978).
- 991 36. Stone, L. S. & Lisberger, S. G. Visual responses of Purkinje cells in the cerebellar flocculus
992 during smooth-pursuit eye movements in monkeys. I. Simple spikes. *J. Neurophysiol.* **63**,
993 1241–1261 (1990).
- 994 37. Krauzlis, R. J. & Lisberger, S. G. Directional organization of eye movement and visual
995 signals in the floccular lobe of the monkey cerebellum. *Exp. Brain Res.* **109**, 289–302
996 (1996).
- 997 38. Lisberger, S. G. & Fuchs, A. F. Role of primate flocculus during rapid behavioral
998 modification of vestibuloocular reflex. II. Mossy fiber firing patterns during horizontal head
999 rotation and eye movement. *J. Neurophysiol.* **41**, 764–777 (1978).
- 1000 39. Buonomano, D. V. & Mauk, M. D. Neural Network Model of the Cerebellum: Temporal
1001 Discrimination and the Timing of Motor Responses. *Neural Comput.* **6**, 38–55 (1994).

- 1002 40. Medina, J. F., Garcia, K. S., Nores, W. L., Taylor, N. M. & Mauk, M. D. Timing
1003 Mechanisms in the Cerebellum: Testing Predictions of a Large-Scale Computer Simulation.
1004 *J. Neurosci.* **20**, 5516–5525 (2000).
- 1005 41. Yamazaki, T. & Tanaka, S. Computational Models of Timing Mechanisms in the Cerebellar
1006 Granular Layer. *Cerebellum Lond. Engl.* **8**, 423–432 (2009).
- 1007 42. Gilmer, J. I. *et al.* An emergent temporal basis set robustly supports cerebellar time-series
1008 learning. *J. Neurophysiol.* **129**, 159–176 (2023).
- 1009 43. Marr, D. A theory of cerebellar cortex. *J. Physiol.* **202**, 437-470.1 (1969).
- 1010 44. Albus, J. S., Branch, D. T., Donald, C. & Perkel, H. *A Theory of Cerebellar Function.*
1011 (1971).
- 1012 45. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–56
1013 (2012).
- 1014 46. Medina, J. F. & Lisberger, S. G. Links from complex spikes to local plasticity and motor
1015 learning in the cerebellum of awake-behaving monkeys. *Nat. Neurosci.* **11**, 1185–1192
1016 (2008).
- 1017 47. Yang, Y. & Lisberger, S. G. Interaction of plasticity and circuit organization during the
1018 acquisition of cerebellum-dependent motor learning. *eLife* **2**, e01574 (2013).
- 1019 48. Fleming, E. A., Field, G. D., Tadross, M. R. & Hull, C. Local synaptic inhibition mediates
1020 cerebellar granule cell pattern separation and enables learned sensorimotor associations. *Nat.*
1021 *Neurosci.* 1–13 (2024) doi:10.1038/s41593-023-01565-4.
- 1022 49. D'Angelo, E. *et al.* The cerebellar Golgi cell and spatiotemporal organization of granular
1023 layer activity. *Front. Neural Circuits* **7**, 93 (2013).
- 1024 50. D'Angelo, E. The Critical Role of Golgi Cells in Regulating Spatio-Temporal Integration
1025 and Plasticity at the Cerebellum Input Stage. *Front. Neurosci.* **2**, 35–46 (2008).
- 1026 51. Holt, G. R., Softky, W. R., Koch, C. & Douglas, R. J. Comparison of discharge variability in
1027 vitro and in vivo in cat visual cortex neurons. *J. Neurophysiol.* **75**, 1806–1814 (1996).
- 1028 52. Kim, J. & Augustine, G. J. Molecular Layer Interneurons: Key Elements of Cerebellar
1029 Network Computation and Behavior. *Neuroscience* **462**, 22–35 (2021).
- 1030 53. Mugnaini, E., Sekerková, G. & Martina, M. The unipolar brush cell: A remarkable neuron
1031 finally receiving deserved attention. *Brain Res. Rev.* **66**, 220–245 (2011).
- 1032 54. Balmer, T. S. & Trussell, L. O. Selective targeting of unipolar brush cell subtypes by
1033 cerebellar mossy fibers. *eLife* **8**, e44964 (2019).
- 1034 55. Guo, C., Huson, V., Macosko, E. Z. & Regehr, W. G. Graded heterogeneity of metabotropic
1035 signaling underlies a continuum of cell-intrinsic temporal responses in unipolar brush cells.
1036 *Nat. Commun.* **12**, 5491 (2021).
- 1037 56. Huson, V. & Regehr, W. G. Realistic mossy fiber input patterns to unipolar brush cells evoke
1038 a continuum of temporal responses comprised of components mediated by different
1039 glutamate receptors. *eLife* **13**, (2024).
- 1040 57. Barmack, N. H. & Yakhnitsa, V. Functions of interneurons in mouse cerebellum. *J.*
1041 *Neurosci.* **28**, 1140–1152 (2008).
- 1042 58. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation
1043 by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
- 1044 59. Rajalingham, R., Piccato, A. & Jazayeri, M. Recurrent neural networks with explicit
1045 representation of dynamic latent variables can mimic behavioral patterns in a physical
1046 inference task. *Nat. Commun.* **13**, 5865 (2022).

- 1047 60. Sussillo, D., Churchland, M. M., Kaufman, M. T. & Shenoy, K. V. A neural network that
1048 finds a naturalistic solution for the production of muscle activity. *Nat. Neurosci.* **18**, 1025–
1049 1033 (2015).
- 1050 61. Palkovits, M., Magyar, P. & Szentágothai, J. Quantitative histological analysis of the
1051 cerebellar cortex in the cat: II. Cell numbers and densities in the granular layer. *Brain Res.*
1052 **32**, 15–30 (1971).
- 1053 62. Jakab, R. L. & Hámori, J. Quantitative morphology and synaptology of cerebellar glomeruli
1054 in the rat. *Anat. Embryol. (Berl.)* **179**, 81–88 (1988).
- 1055 63. Gibbon, J. Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* **84**,
1056 279–325 (1977).
- 1057 64. Medina, J. F. & Mauk, M. D. Computer simulation of cerebellar information processing.
1058 *Nat. Neurosci.* **3**, 1205–1211 (2000).
- 1059 65. Bullock, D., Fiala, J. C. & Grossberg, S. A neural model of timed response learning in the
1060 cerebellum. *Neural Netw.* **7**, 1101–1114 (1994).
- 1061 66. Barri, A., Wiechert, M. T., Jazayeri, M. & DiGregorio, D. A. Synaptic basis of a sub-second
1062 representation of time in a neural circuit model. *Nat. Commun.* **13**, 7902 (2022).
- 1063 67. Ahmadian, Y. & Miller, K. D. What is the dynamical regime of cerebral cortex? *Neuron* **109**,
1064 3373–3391 (2021).
- 1065 68. Xue, M., Atallah, B. V. & Scanziani, M. Equalizing excitation–inhibition ratios across visual
1066 cortical neurons. *Nature* **511**, 596–600 (2014).
- 1067 69. Shu, Y., Hasenstaub, A. & McCormick, D. A. Turning on and off recurrent balanced cortical
1068 activity. *Nature* **423**, 288–293 (2003).
- 1069 70. Wehr, M. & Zador, A. M. Balanced inhibition underlies tuning and sharpens spike timing in
1070 auditory cortex. *Nature* **426**, 442–446 (2003).
- 1071 71. Medina, J. F., Carey, M. R. & Lisberger, S. G. The Representation of Time for Motor
1072 Learning. *Neuron* **45**, 157–167 (2005).
- 1073 72. Herzfeld, D. J., Hall, N. J., Tringides, M. & Lisberger, S. G. Principles of operation of a
1074 cerebellar learning circuit. *eLife* **9**, e55217 (2020).
- 1075 73. Yang, Y. & Lisberger, S. G. Purkinje-cell plasticity and cerebellar motor learning are graded
1076 by complex-spike duration. *Nature* **510**, 529–532 (2014).
- 1077 74. Hall, N. J., Yang, Y. & Lisberger, S. G. Multiple components in direction learning in smooth
1078 pursuit eye movements of monkeys. *J. Neurophysiol.* (2018) doi:10.1152/jn.00261.2018.
- 1079 75. Yang, Y. & Lisberger, S. G. Role of Plasticity at Different Sites across the Time Course of
1080 Cerebellar Motor Learning. *J. Neurosci.* **34**, 7077–7090 (2014).
- 1081 76. Badura, A. *et al.* Climbing Fiber Input Shapes Reciprocity of Purkinje Cell Firing. *Neuron*
1082 **78**, 700–713 (2013).
- 1083 77. Stone, L. S. & Lisberger, S. G. Visual responses of Purkinje cells in the cerebellar flocculus
1084 during smooth-pursuit eye movements in monkeys. II. Complex spikes. *J. Neurophysiol.* **63**,
1085 1262–1275 (1990).
- 1086 78. Larry, N., Yarkoni, M., Lixenberg, A. & Joshua, M. Cerebellar climbing fibers encode
1087 expected reward size. *eLife* **8**, e46870 (2019).
- 1088 79. Hopfield, J. J. Transforming neural computations and representing time. *Proc. Natl. Acad.*
1089 *Sci.* **93**, 15440–15444 (1996).
- 1090 80. Hopfield, J. J. & Tank, D. W. Computing with Neural Circuits: A Model. *Science* **233**, 625–
1091 633 (1986).
- 1092 81. Luo, L. Architectures of neuronal circuits. *Science* **373**, eabg7285 (2021).

- 1093 82. Mathis, M. W., Perez Rotondo, A., Chang, E. F., Tolias, A. S. & Mathis, A. Decoding the
1094 brain: From neural representations to mechanistic models. *Cell* **187**, 5814–5832 (2024).
- 1095 83. Luo, L., Callaway, E. M. & Svoboda, K. Genetic Dissection of Neural Circuits. *Neuron* **57**,
1096 634–660 (2008).
- 1097 84. Fishell, G. & Heintz, N. The Neuron Identity Problem: Form Meets Function. *Neuron* **80**,
1098 602–612 (2013).
- 1099 85. Ecker, J. R. *et al.* The BRAIN Initiative Cell Census Consortium: Lessons Learned toward
1100 Generating a Comprehensive Brain Cell Atlas. *Neuron* **96**, 542–557 (2017).
- 1101 86. Zeng, H. & Sanes, J. R. Neuronal cell-type classification: challenges, opportunities and the
1102 path forward. *Nat. Rev. Neurosci.* **18**, 530–546 (2017).
- 1103 87. Migliore, M. & Shepherd, G. M. An integrated approach to classifying neuronal phenotypes.
1104 *Nat. Rev. Neurosci.* **6**, 810–818 (2005).
- 1105 88. Gouwens, N. W. *et al.* Classification of electrophysiological and morphological neuron types
1106 in the mouse visual cortex. *Nat. Neurosci.* **22**, 1182–1195 (2019).
- 1107 89. Poulin, J.-F., Tasic, B., Hjerling-Leffler, J., Trimarchi, J. M. & Awanramani, R.
1108 Disentangling neural cell diversity using single-cell transcriptomics. *Nat. Neurosci.* **19**,
1109 1131–1141 (2016).
- 1110 90. Josh Huang, Z. & Zeng, H. Genetic Approaches to Neural Circuits in the Mouse. *Annu. Rev.*
1111 *Neurosci.* **36**, 183–215 (2013).
- 1112 91. Gordon, N. Speech, language, and the cerebellum. *Eur. J. Disord. Commun. J. Coll. Speech*
1113 *Lang. Ther. Lond.* **31**, 359–367 (1996).
- 1114 92. Schmahmann, J. D. & Sherman, J. C. The cerebellar cognitive affective syndrome. *Brain*
1115 **121**, 561–579 (1998).
- 1116 93. Van Overwalle, F. Social and emotional learning in the cerebellum. *Nat. Rev. Neurosci.* **25**,
1117 776–791 (2024).
- 1118 94. Schmahmann, J. D. The cerebellum and cognition. *Neurosci. Lett.* **688**, 62–75 (2019).
- 1119 95. Straub, I. *et al.* Gradients in the mammalian cerebellar cortex enable Fourier-like
1120 transformation and improve storing capacity. *eLife* **9**, e51771 (2020).
- 1121 96. Masoli, S., Tognolina, M., Laforenza, U., Moccia, F. & D’Angelo, E. Parameter tuning
1122 differentiates granule cell subtypes enriching transmission properties at the cerebellum input
1123 stage. *Commun. Biol.* **3**, 1–12 (2020).
- 1124 97. Diedrichsen, J., King, M., Hernandez-Castillo, C., Sereno, M. & Ivry, R. B. Universal
1125 Transform or Multiple Functionality? Understanding the Contribution of the Human
1126 Cerebellum across Task Domains. *Neuron* **102**, 918–928 (2019).
- 1127 98. Schmahmann, J. D. From movement to thought: Anatomic substrates of the cerebellar
1128 contribution to cognitive processing. *Hum. Brain Mapp.* **4**, 174–198 (1996).
- 1129 99. Heiney, S. A., Wohl, M. P., Chettih, S. N., Ruffolo, L. I. & Medina, J. F. Cerebellar-
1130 Dependent Expression of Motor Learning during Eyeblink Conditioning in Head-Fixed
1131 Mice. *J. Neurosci.* **34**, 14845–14853 (2014).
- 1132 100. Ohayama, T. & Mauk, M. D. Cerebellar learning. in *Neurobiology of learning and*
1133 *memory*, 2nd ed 427–456 (Elsevier Academic Press, San Diego, CA, US, 2007).
1134 doi:10.1016/B978-012372540-0/50014-5.
- 1135 101. Hansel, C., Linden, D. J. & D’Angelo, E. Beyond parallel fiber LTD: the diversity of
1136 synaptic and non-synaptic plasticity in the cerebellum. *Nat. Neurosci.* **4**, 467–475 (2001).
- 1137 102. Tabata, T. & Kano, M. Synaptic Plasticity in the Cerebellum. in *Handbook of*
1138 *Neurochemistry and Molecular Neurobiology: Neural Signaling Mechanisms* (eds. Lajtha,

- 1139 A. & Mikoshiba, K.) 63–86 (Springer US, Boston, MA, 2009). doi:10.1007/978-0-387-
1140 30370-3_6.
- 1141 103. Suvrathan, A., Payne, H. L. & Raymond, J. L. Timing Rules for Synaptic Plasticity
1142 Matched to Behavioral Function. *Neuron* **92**, 959–967 (2016).
- 1143 104. Payne, H. L. *et al.* Cerebellar Purkinje cells control eye movements with a rapid rate code
1144 that is invariant to spike irregularity. *eLife* **8**, e37102 (2019).
- 1145 105. Lisberger, S. G. & Pavelko, T. A. Brain stem neurons in modified pathways for motor
1146 learning in the primate vestibulo-ocular reflex. *Science* **242**, 771–773 (1988).
- 1147 106. Lisberger, S. G. Visual Guidance of Smooth Pursuit Eye Movements. *Annu. Rev. Vis. Sci.*
1148 **1**, 447–468 (2015).
- 1149 107. Krauzlis, R. J. & Lisberger, S. G. Simple spike responses of gaze velocity Purkinje cells
1150 in the floccular lobe of the monkey during the onset and offset of pursuit eye movements. *J.
1151 Neurophysiol.* **72**, 2045–2050 (1994).
- 1152 108. Crowley, J. J., Fioravante, D. & Regehr, W. G. Dynamics of Fast and Slow Inhibition
1153 from Cerebellar Golgi Cells Allow Flexible Control of Synaptic Integration. *Neuron* **63**,
1154 843–853 (2009).
- 1155 109. Kennedy, A. *et al.* A temporal basis for predicting the sensory consequences of motor
1156 commands in an electric fish. *Nat. Neurosci.* **17**, 416–422 (2014).
- 1157 110. Zampini, V. *et al.* Mechanisms and functional roles of glutamatergic synapse diversity in
1158 a cerebellar circuit. *eLife* **5**, e15872 (2016).
- 1159 111. Leigh, R. J. & Zee, D. S. *The Neurology of Eye Movements*. (Oxford University Press,
1160 2015).
- 1161 112. Ramachandran, R. & Lisberger, S. G. Normal Performance and Expression of Learning
1162 in the Vestibulo-Ocular Reflex (VOR) at High Frequencies. *J. Neurophysiol.* **93**, 2028–2038
1163 (2005).
- 1164 113. Robinson, D. A. A Method of Measuring Eye Movement Using a Scleral Search Coil in a
1165 Magnetic Field. *IEEE Trans. Bio-Med. Electron.* **10**, 137–145 (1963).
- 1166 114. Rashbass, C. The relationship between saccadic and smooth tracking eye movements. *J.
1167 Physiol.* **159**, 326–338 (1961).
- 1168 115. de Brouwer, S., Yuksel, D., Blohm, G., Missal, M. & Lefèvre, P. What Triggers Catch-
1169 Up Saccades During Visual Tracking? *J. Neurophysiol.* **87**, 1646–1650 (2002).
- 1170 116. Hall, N. J., Herzfeld, D. J. & Lisberger, S. G. Evaluation and resolution of many
1171 challenges of neural spike sorting: a new sorter. *J. Neurophysiol.* **126**, 2065–2090 (2021).
- 1172 117. Herzfeld, D. J. & Beardsley, S. A. Synaptic weighting for physiological responses in
1173 recurrent spiking neural networks. in *2011 Annual International Conference of the IEEE
1174 Engineering in Medicine and Biology Society, EMBC* 4187–4190 (2011).
1175 doi:10.1109/IEMBS.2011.6091039.
- 1176 118. Ruigrok, T. J. H., Hensbroek, R. A. & Simpson, J. I. Spontaneous Activity Signatures of
1177 Morphologically Identified Interneurons in the Vestibulocerebellum. *J. Neurosci.* **31**, 712–
1178 724 (2011).
- 1179 119. Walsh, J. V., Houk, J. C., Atluri, R. L. & Mugnaini, E. Synaptic Transmission at Single
1180 Glomeruli in the Turtle Cerebellum. *Science* **178**, 881–883 (1972).
- 1181 120. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. Preprint at
1182 <https://doi.org/10.48550/arXiv.1412.6980> (2017).
- 1183 121. Loshchilov, I. & Hutter, F. SGDR: Stochastic Gradient Descent with Warm Restarts.
1184 Preprint at <https://doi.org/10.48550/arXiv.1608.03983> (2017).

1185 122. Herzfeld, D. J., Kojima, Y., Soetedjo, R. & Shadmehr, R. Encoding of error and learning
1186 to correct that error by the Purkinje cells of the cerebellum. *Nat. Neurosci.* **21**, 736–743
1187 (2018).
1188