Synaptic dynamics are a tunable substrate sculpting neural population activity

Franziska Bender[1] *, B. Semihcan Sermet[1] *, Stefano Borda Bossana[1], Alessandro Barri[1], Selin Schamiloglu[2], Giovanni Diana[1], Maria-Miruna Costreie[1], Gael Moneron[1], Adam W. Hantman[3,4], and David A. DiGregorio[1,2]

Affiliations:

[1]Institut Pasteur, Université Paris Cité, CNRS UMR 3571, Synapse and Circuit Dynamics Laboratory, Paris, France.
[2]Department of Physiology and Biophysics, University of Colorado School of Medicine, Aurora, Colorado, USA 80045.
[3]Department of Cell Biology and Physiology, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.
[4]UNC Neuroscience Center, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

*co-first authors

Correspondence: david.digregorio@cuanschutz.edu

Abstract:

A central tenet of cerebellar computation posits that granule cells generate sparse spatiotemporal activity patterns that support precisely timed motor and cognitive outputs. Using high-speed *in vivo* calcium and glutamate imaging combined with slice electrophysiology, we show that heterogeneous synaptic dynamics transform mossy fiber inputs into temporally sparse, sequential patterns of GC activity. Region-specific differences in MF glutamate release shape GC response duration and sparsity, tuning the temporal statistics of population sequences to match sensorimotor associative learning demands. These findings provide direct evidence for temporally sparse GC sequences and establish synaptic dynamics as a ubiquitous, tunable mechanism structuring neural activity in time.

Introduction

Neural population activity across the brain unfolds as temporally structured sequences that underlie perception, memory, and precisely timed motor behaviors (*1–14*). A central organizing principle of these representations is sparsity: restricting activity to a fraction of neurons enhances storage capacity, robustness, and separability of neural codes (*2*, *15–20*). In the cerebellum, the seminal theories of David Marr and James Albus proposed that expansion of mossy fiber (MF) inputs onto a vastly larger population of granule cells (GCs) generates sparse, high-dimensional representations for pattern separation (*16*, *21–23*). This framework was later extended to the temporal domain, proposing that diverse GC response patterns form a high-dimensional temporal feature space that can be linearly decoded by Purkinje cells to produce precisely timed outputs (*24–27*).

Despite the decades of influence, direct experimental evidence of these theories has been lacking, in part due to limitations in the temporal resolution of typical calcium indicators (*28*) and the inability to resolve GC units with high-density electrophysiology (*29*). In particular, it remains unknown whether GC activity is temporally sparse and additionally contributes to pattern separation (*30*, *31*). Work in other brain regions shows that temporally sparse population responses that tile time form an efficient substrate for downstream linear decoding (*7*, *9*), and that sequence timescales are tuned to computational demands of different brain regions (*7*). However, whether such principles apply to the cerebellar GC layer remains unresolved.

Despite the ubiquity of neural sequences throughout the brain (*14*), a major outstanding question is therefore what circuit mechanisms generate temporally sparse neural sequences. Proposed explanations include structured connectivity, recurrent dynamics, and intrinsic cellular heterogeneity (*2*, *7*, *32–34*). Because short-term synaptic plasticity (STP) is ubiquitous across the nervous system (*35*) and easy to modulate (*36*), diversity in STP represents a compelling and general mechanism for shaping temporally structured population activity and its regional specialization. We therefore explored whether the known diversity of MF-GC synapses could be a mechanism to control the temporal statistics (sparsity and tiling) of GC population responses.

Here, we combine ex vivo brain slice recordings with high-speed *in vivo* two-photon imaging using ultrafast calcium and glutamate sensors to show that regional diversity in MF–GC STP provides a synaptic mechanism that converts dense afferent activity into temporally sparse GC sequences. By tuning synaptic timescales, this mechanism sets the temporal structure of GC activity to match the computational demands of distinct cerebellar regions.

Results

GC neural sequences form temporally sparse sensory representations

To determine whether cerebellar granule cells (GCs) form a temporal basis set resembling a neuronal activity sequence, we performed high-speed two-photon imaging of intracellular calcium responses evoked by air puff stimulation of the mouse whisker field (1-s air puff at 5 psi; **Fig. 1A**, **fig. S1A**). Resolving GC population dynamics at sub-second resolution *in vivo* is challenging, as GC spiking is inaccessible to modern silicon probes (*29*), and older genetically encoded calcium indicators blur rapid temporal structure (*37*–*39*). We therefore used the fast calcium indicator GCaMP8f (*40*) to resolve stimulus-locked GC population activity in Crus I. Whisker–air-puff–evoked responses were elicited using stimulus parameters similar to those used as conditioned stimuli in timing tasks (*41*), for which GC population dynamics have been hypothesized to provide a temporal basis for downstream learning (*27*). Whisker air puff produced calcium responses in ~2% of identified GCs (**fig. S1B**), consistent with theoretical predictions that GCs form a sparse population representation of sensory-motor features (*16*, *42*, *43*).

We next examined the temporal characteristics of calcium responses within this sparsely recruited GC ensemble. GC responses were transient and generally shorter than the stimulus duration. They reliably produced peak response times that tiled the stimulus duration and were consistent across trials (**fig. S1**), reminiscent of a neural sequence. Individual GC responses, both within a single field of view and across animals, occurred as discrete calcium responses. Onset times spanned nearly an order of magnitude (10 ms–0.9 s). Response durations were heterogeneous (full-width-at-half-maximum (FWHM): 0.05–3.8 s; **Fig. 1**). At the population level, the temporal representation was biased toward GCs generating brief calcium transients near whisker air puff onset, whereas responses peaking later in the stimulus were progressively broader and less numerous. Consistent with this structure, peak response latency was positively correlated with burst duration (**fig. S2A**). We also observed a similar sequence-like population activity in GC axons within the molecular layer (**fig. S2B-D**). This confirmed that temporally sparse activity sequences found in the GC layer are also present across the population of GC axons likely to synapse onto a Purkinje cell (PC). In contrast, sequence-like population activity was not observed in molecular layer interneurons (MLIs), which exhibited shorter onset times and times-to-peak (**fig. S2E, F**). This suggests that sequence-like population activity is specific to GCs. Overall, results demonstrate that GC population activity is organized in a sequence-like manner

114 at the sub-second time scale and suggest that diverse GC population activity can serve as a
115 temporal basis set, as proposed by longstanding theories of cerebellar learning (*27*).
116
117 **GC neural sequence temporal sparsity requires a local transformation of MF input**
118 The observed GC sequences could be inherited from their mossy fiber (MF) inputs or
119 instead arise from a transformation of those inputs. To distinguish between these possibilities, we
120 monitored MF activity projecting from the whisker somatosensory and motor cortices through the
121 basal pons (secondary somatosensory inputs (*44*)) or directly via the trigeminal nucleus principalis
122 (Pr5; primary somatosensory inputs, **Fig. 1B,C, fig. S3**) (*45–47*) to Crus I using the same calcium
123 indicator and whisker stimulus. The fraction of activated MFs ranged from 16% in pontine MFs to
124 32% in Pr5 MFs (**fig. S1B**), clearly denser than the GC population (2%). Calcium responses were
125 mostly excitatory, but we also observed inhibitory responses (Crus I GCs: 29% of responsive cells,
126 Crus I pontine MFs: 5%, Crus I Pr5 MFs: 4%) and a small fraction of "off" responses, except in
127 Pr5 MFs (Crus I GCs: 2%, Crus I pontine MFs: 1%, **fig. S4**). We focused subsequent analysis on
128 excitatory responses.
129 The time course of MF responses was consistently slower than GC responses, exhibiting
130 broader FWHM values (**Fig. 1D, left**) corresponding to a lower temporal sparsity, as quantified
131 using a lifetime sparsity metric (**Fig. 1D, right**) (*48*). Cross-validation of MF and GC calcium
132 responses to whisker air puffs demonstrated that responses in both populations were robust and
133 reliable (**fig. S1**). The kinetic differences were not attributable to the calcium indicator, as *ex vivo*
134 recordings showed that single- and burst-evoked calcium responses in MF terminals and GC
135 somata had comparable rise and decay times (**fig. S5**). The FWHM of MF or GC calcium
136 responses did not correlate with whisker air puff response duration (**fig. S6A-C**). Hence, GC
137 population activity primarily reflected a sparse temporal code of the whisker air puff rather than
138 motor-related responses. No correlation was observed between maximal deflection and response
139 duration in GCs, Pr5 MFs, or pontine MFs (**fig. S6C**), further supporting a predominantly sensory
140 representation. Analysis of the response onset variability showed that GCs exhibit similar
141 heterogeneity (coefficient of variation, CV = 1.5) to that of their pontine MFs (CV = 1.6, two-sample
142 Kolmogorov–Smirnov test, D = 0.119, $P = 0.43$). Pr5 MF response onsets, in contrast, were earlier
143 and less diverse than GCs (CV = 1.0, two-sample Kolmogorov–Smirnov test, D = 0.43, $P < 0.0001$,
144 **fig. S6D**). The short response onset of Pr5 MFs and the heterogeneity in pontine MF onset times
145 (**fig. S6D**) are consistent with the direct arrival of the sensory information from the trigeminal
146 nucleus and the delayed cortical information arising from cortico-pontine projections (*49*, *50*),
147 respectively. Together, these results indicate that the GC layer itself performs a critical

transformation, converting dense MF temporal representations of sensory input into a temporally sparse sequence of GC activity that can serve as a basis set from which precise output activity dynamics can be learned (*7*, *9*).

Region-specific timescales of GC sequences

Cerebellar circuits support behaviors with distinct temporal requirements, and synaptic plasticity rules are tuned across cerebellar regions to match these timing demands (*51*, *52*), raising the possibility that MF–GC transformations are region-specific and shape the temporal structure of local neural sequences. To determine whether MF–GC transformations generate distinct region-specific GC sequences, we measured Simplex GC and MF activity in response to whisker air puff (**Fig. 1E-G**). We took advantage of the fact that Crus I and Simplex regions both receive projections from Pr5 and the whisker somatosensory cortex (through the BPN, **fig. S3**) (*45*). No difference was observed between Crus I and Simplex in the relative number of activated GCs (3% in Simplex, similar to 2% in Crus I, **fig. S1B**). When comparing the temporal statistics of GC sequences, we observed no differences in the response width or lifetime sparsity of MFs between Crus I and Simplex (**Fig. 1D**). In contrast, GC responses from Crus I were significantly shorter and had higher lifetime sparsity than those from Simplex (**Fig. 1D**). Thus, despite receiving MF inputs with similar temporal statistics, MF-GC transformations are region-specific, generating GC activity with greater temporal sparsity in Crus I than in Simplex.

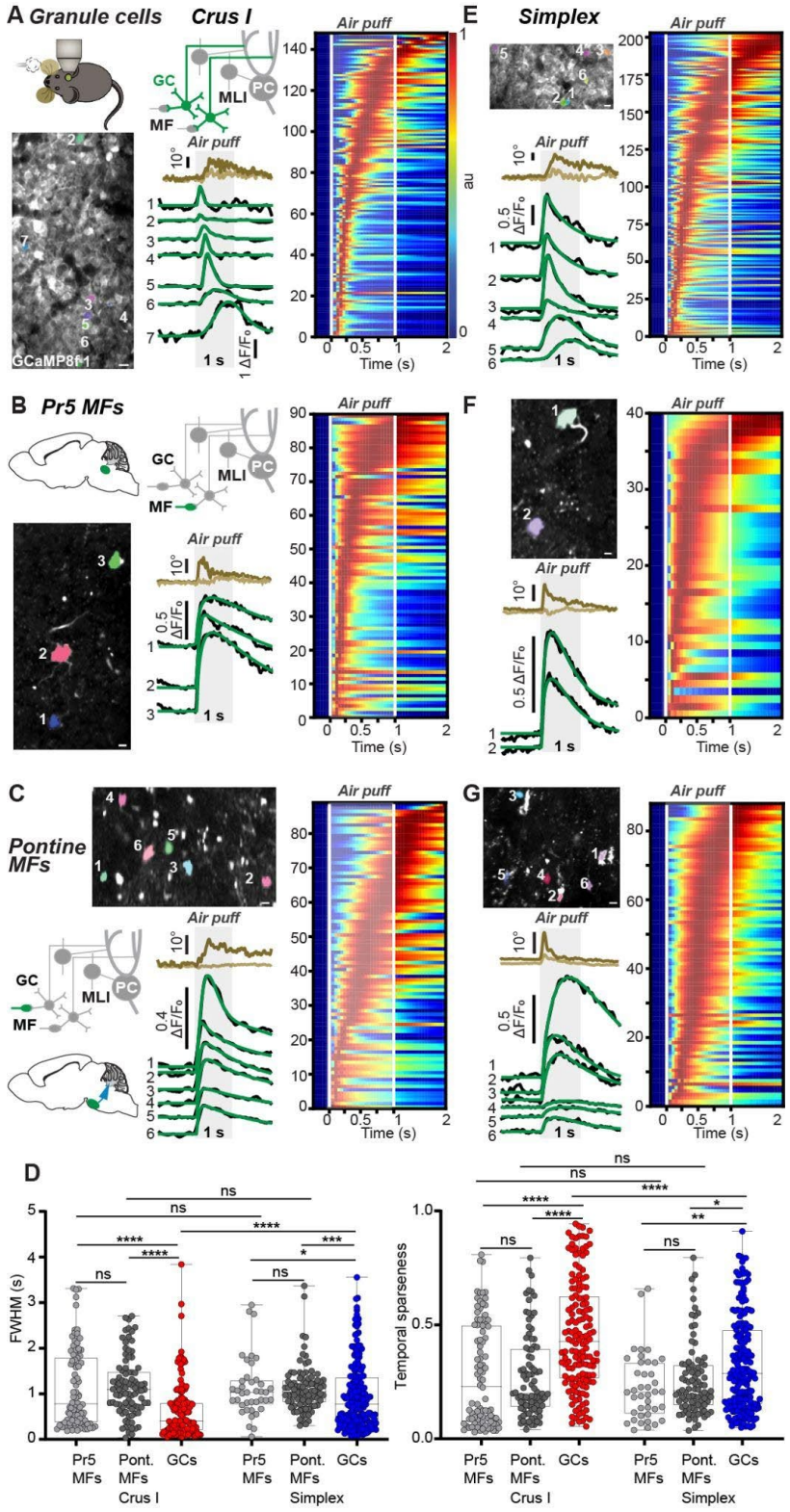
A Granule cells
Crus I
GC
MLI
PC
MF
Air puff
GCaMP8f-1
1 s
1 ΔF/Fo
Time (s)
au
E
Simplex
B Pr5 MFs
F
C
Pontine MFs
G
D
FWHM (s)
Temporal sparseness
Pr5 MFs
Pont. MFs
GCs
Crus I
Simplex
ns
****
***
**
*

Fig. 1. Crus I and Simplex mossy fiber and granule cell temporal representations of whisker air puffs. (A) Left, two-photon imaging of Crus I granule cells (GCs) expressing GCaMP8f. Center, whisker-air-puff responses of seven simultaneously recorded GCs averaged over 30 trials; the 1-s air puff is indicated by grey shading. Responses are shown above (light brown, ipsilateral; dark brown, contralateral to air puff). Right, average Crus I GC responses across experiments (N = 10 mice, n = 148 cells) shown as raster plots, normalized and sorted by time to peak. MLI: Molecular layer interneuron. PC: Purkinje cell. (**B and C**) Crus I mossy fiber (MF) responses to whisker air puffs from Pr5 (B; N = 4 mice, n = 90 cells) and pontine inputs (C; N = 3 mice, n = 89 cells). **(D)** Left, in Crus I and Simplex, response duration (full width at half maximum, FWHM) was greater in Pr5 (1.12 ± 0.09 s, $P < 0.0001$; Pr5 MFs, 1.13 ± 0.10 s, $P = 0.027$) and pontine MFs (1.16 ± 0.07 s, $P < 0.0001$; 1.12 ± 0.06 s, $P = 0.0003$) than in GCs (0.62 ± 0.05 s; GCs, 0.95 ± 0.05 s), with no difference between MF pathways ($P = 0.17$; $P = 0.86$). Crus I GC responses were shorter than Simplex GC responses ($P < 0.0001$), whereas Pr5 and pontine MF responses did not differ between regions (P = 0.80 and $P = 0.34$, respectively). Right, GC responses were temporally sparser than MF responses in Crus I (GCs, 0.46 ± 0.02; pontine MFs, 0.27 ± 0.02, $P < 0.0001$; Pr5 MFs, 0.28 ± 0.02, $P < 0.0001$) and Simplex (GCs, 0.33 ± 0.01; pontine MFs, 0.29 ± 0.02, $P = 0.017$; Pr5 MFs, 0.23 ± 0.02, $P = 0.004$). No difference was observed between Pr5 MFs and pontine MFs in Crus I ($P = 0.1$) or Simplex ($P = 0.09$). GC responses were more temporally sparse in Crus I than in Simplex ($P < 0.0001$), whereas no difference between regions was observed in MFs (Pr5, $P = 0.86$; pontine, $P = 0.72$). (**E–G**) Simplex responses to whisker air puffs: GCs (**E**; N = 8 mice, n = 203 cells), Pr5 MFs (**F**; N = 3 mice, n = 40 cells), and pontine MFs (**G**; N = 3 mice, n = 88 cells). Scale bars, 5 µm. Data are shown as median ± interquartile range, values in text as mean ± SEM. Statistical significance: $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***), $P < 0.0001$ (****).

Heterogeneity of Crus I and Simplex MF-GC synaptic dynamics in acute brain slices

To investigate what may underlie the difference in temporal sparsity between Crus I and Simplex, we next directly tested whether MF-GC synapses exhibit diverse and region-specific short-term plasticity (STP, **Fig. 1**). While cerebellar cortical models incorporating heterogeneous MF–GC STP predict that step changes in MF firing can be transformed into diverse GC temporal basis sets (*53*), evidence for regional diversity is lacking. We therefore recorded unitary MF–GC AMPA receptor–mediated excitatory postsynaptic currents (EPSCs) in whole-cell voltage-clamp in acute cerebellar slices using a blind minimum-stimulation approach (*54*). EPSCs from both Crus I and Simplex were highly heterogeneous in amplitude, coefficient of variation (CV), and paired-pulse ratio (PPR) across connections (**Fig. 2A, B, fig. S7**), as previously reported in the vestibulo-cerebellum (*31*). Our previous theoretical work showed that the time course of MF–GC synaptic dynamics during presynaptic trains directly determines the temporal statistics of GC firing responses, including their duration (*53*). We therefore quantified the timescales of STP at MF–GC synapses in Crus I and Simplex by estimating a weighted decay time constant (τ_{syn}) associated with the adaptation of EPSC amplitudes during prolonged presynaptic firing (*53*) (**Fig. 2C**, **Materials and Methods**). Across three stimulation frequencies, τ_{syn} was faster in Crus I than in Simplex (**Fig. 2D, E**). Thus, diversity in the STP timescale provides a candidate mechanism for the MF-GC transformation and the differences in GC temporal sparsity we observe between Crus I and Simplex.

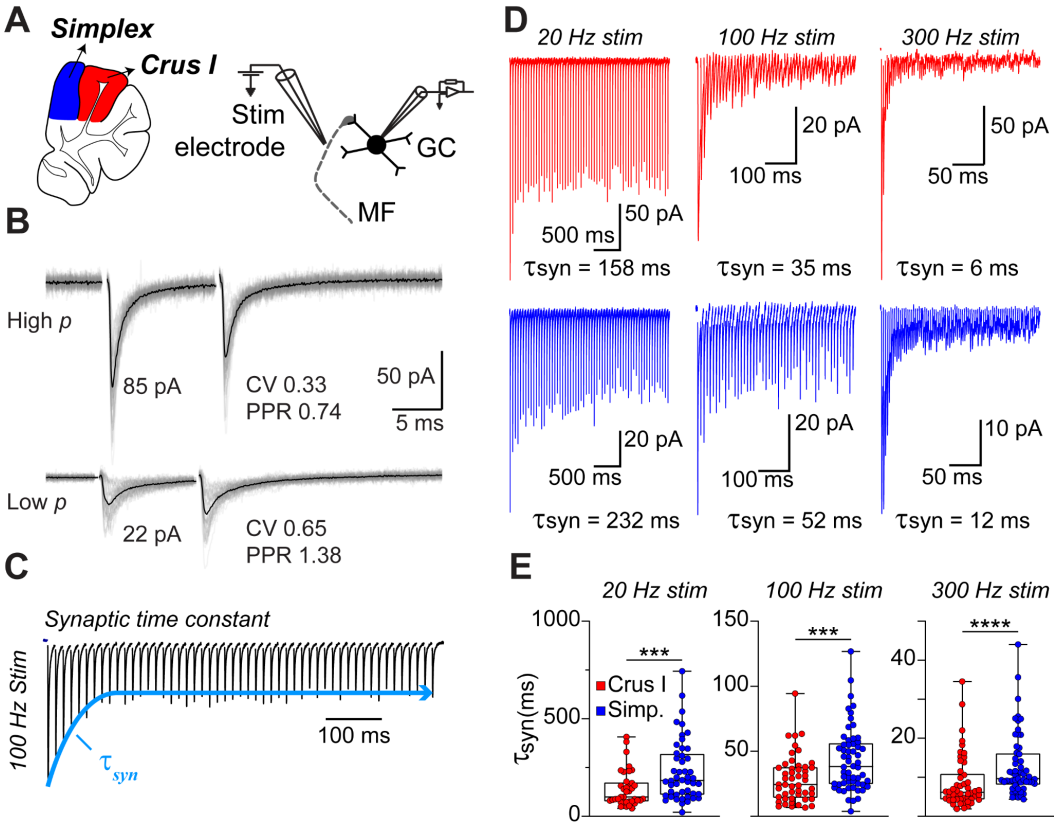

Fig. 2. Regional differences in short-term synaptic dynamics at MF-GC synapses. (A) Schematic demonstrating the experimental design of whole-cell recordings from GCs in Crus I (red) or Simplex (blue) during MF stimulation in acute brain slices. **(B)** Unitary EPSCs from individual trials (gray traces) and mean (black trace) from two example MF-GC synapses with high (top) and low (bottom) release probability. EPSC peak amplitude, coefficient of variation of the peak amplitude (CV), and paired-pulse ratio (PPR) were measured for each synapse from responses to 100 Hz MF stimulation (see **fig. S7**). **(C)** Schematic illustrating exponential fit to EPSC peak amplitudes to extract synaptic time constants (τ_{syn}). **(D)** Example EPSC traces (artifact subtracted and baseline corrected (see **Materials and Methods**) evoked by trains of MF stimulation at 20, 100, and 300 Hz in Crus I (top, red) and Simplex (bottom, blue). **(E)** Synaptic time constants (τ_{syn}) at 20, 100, and 300 Hz in Crus I (red, n = 51) and Simplex (blue, n = 58). Data are presented as median with interquartile range. Crus I synapses display shorter τ_{syn} than Simplex synapses at 20 Hz (Crus I: 98.85 ± 90.98 ms, Simplex: 183.58 ± 156.72 ms; P = 0.0007, Mann-Whitney test), 100 Hz (Crus I: 24.43 ± 18.04 ms, Simplex: 38.27 ± 24.54 ms; P = 0.0006, Mann-Whitney test), and 300 Hz (Crus I: 6.16 ± 6.81 ms, Simplex: 9.66 ± 8.10 ms; P < 0.0001, Mann-Whitney test) stimulation.

MF–GC synaptic dynamics elicit diverse GC spiking patterns

Although regional differences in MF–GC synaptic dynamics provide a candidate mechanism for GC temporal diversity (*53*), other mechanisms could also contribute to the heterogeneity of whisker-air-puff-evoked GC responses, including Golgi-cell feedback (*27*) and tonic GABAergic inhibition, which influences the threshold for GC firing (*30*, *55*). To isolate the contribution of synaptic dynamics alone, we first tested whether constant-frequency MF stimulation could elicit diverse GC spiking patterns in the absence of GABAergic inhibition. GC action potentials were recorded extracellularly during pharmacological block of both tonic and feedback $GABA_A$ receptor–mediated inhibition (**Fig. 3A**). Despite identical presynaptic drive,

258 individual GCs exhibited remarkably diverse firing dynamics (**Fig. 3B, C**). Most cells responded
259 transiently to sustained 100 Hz MF stimulation, with firing rates that peaked rapidly and decayed
260 with widely varying time constants (~50–800 ms), closely matching predictions from our previous
261 model (*53*) (**Fig. 3C**). Notably, a subset of GCs displayed delayed peak firing (>200 ms),
262 demonstrating that synaptic dynamics alone can generate delayed spike responses, as previously
263 observed for brief stimulus trains (*31*) (**Fig. 3B, magenta, orange**). Thus, intrinsic GC layer
264 mechanisms can transform constant MF input into temporally structured spike output.

265 We next asked whether GC firing diversity directly reflects heterogeneity in MF–GC
266 synaptic dynamics. We reasoned that if synaptic inputs govern spike timing, the time course of
267 EPSC dynamics during high-frequency stimulation (τ_{syn}) should predict the temporal profile of GC
268 spiking (τ_{spike}). To test this, we recorded spike PSTHs from individual GCs in cell-attached mode
269 before transitioning to whole-cell configuration to measure MF–GC EPSCs evoked by identical
270 stimulus trains (**Fig. 3B**). To include GCs with slowly varying firing rates that deviate from
271 exponential decay, we calculated EPSC and PSTH ratios as proxies for τ_{syn} and τ_{spike}. Across
272 cells, we observed a strong correlation between the two (R^2=0.92), establishing a tight synapse-
273 to-spike correspondence (**Fig. 3D**). Consistent with regional differences in MF-GC STP, GC
274 spiking dynamics were significantly slower in Simplex than in Crus I (**Fig. 3E, F),** demonstrating
275 that region-specific synaptic dynamics set the temporal profile of GC output. These results identify
276 STP heterogeneity as a primary determinant of GC firing diversity.

277 To determine if these single-cell mechanisms scale to the population level, we performed
278 high-speed two-photon imaging of GCaMP8f-expressing GCs in acute slices during constant MF
279 stimulation (**Fig. 3G-I**). Identical MF input elicited sequence-like GC population responses with a
280 broad distribution of response latencies and durations (**Fig. 3K, L**). At stimulation frequencies
281 comparable to whisker-air-puff-evoked MF firing in Pr5 (*56*) (100 and 300 Hz), GC responses in
282 Crus I were significantly briefer than those in Simplex (**Fig. 3L, M**), recapitulating the regional
283 differences observed *in vivo* (**Fig. 1**). These population-level dynamics mirror regional
284 specializations in MF–GC STP (**Fig. 2**), supporting the idea that synaptic diversity directly
285 organizes GC temporal sequences. Notably, blocking GABAergic transmission had minimal effect
286 on these temporal patterns (**fig. S8**). Together, these single-cell and population experiments
287 demonstrate that heterogeneity in MF–GC synaptic dynamics is sufficient to generate region-
288 specific, temporally ordered GC firing patterns under constant input, even in the absence of feed-
289 forward and feed-back inhibition.
290

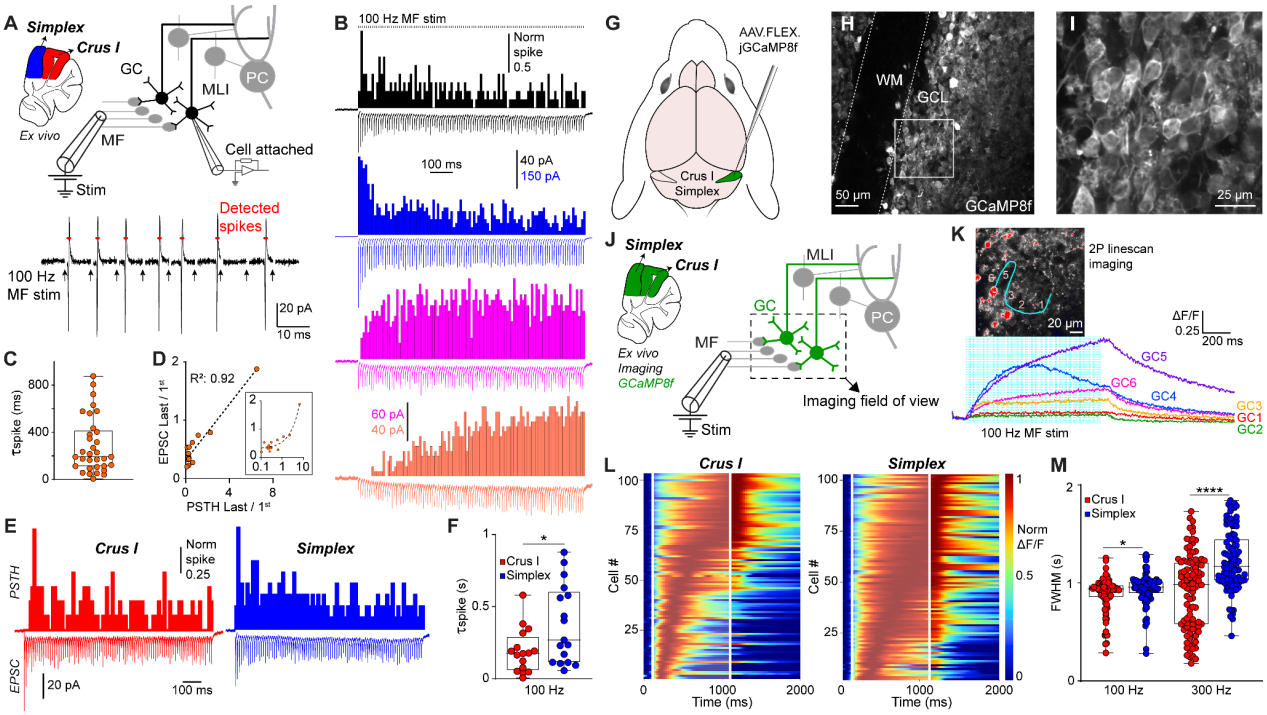

Fig. 3. Synaptic properties shape GC temporal dynamics. (A) Top: Schematic of cell-attached GC recording during MF stimulation. Bottom: Example of 100 Hz MF stimulation and detected GC spikes. Arrows indicate stimulation times; stimulation artefacts were blanked. **(B)** PSTH (above) and mean EPSC (below) responses of four GCs following a 1 s, 100 Hz MF stimulus. Top, black vertical ticks indicate stimulation time points. **(C)** Quantification of the spike decay time constant (τ_{spike}) of all cells responding to 100 Hz MF stimulation (204.81 ± 240.79 ms, n = 33 cells). **(D)** Correlation plot of synaptic EPSC ratio vs spike PSTH ratio (n = 16), measured as the ratio of the steady state EPSC amplitudes or PSTH counts (last) over the first EPSC response amplitude or PSTH bin count. The inset shows the horizontal axis on a log scale (regression line: R^2 = 0.92). **(E)** Representative GC firing PSTHs (top) and average EPSCs (bottom) in Crus I (red, left) and Simplex (blue, right) during 100 Hz MF stimulation. **(F)** Quantification of GC spike decay time constant (τ_{spike}) in Crus I (red, 178.06 ± 156.72 ms, n = 16) and Simplex (blue, 269.14 ± 272.51 ms, n = 17; P = 0.049, Mann-Whitney test). **(G)** Schematic of the AAV injection protocol to express GCaMP8f in Crus I and Simplex GCs. **(H)** Expression of GCaMP8f in GCs in Crus I. WM, white matter; GCL, granule cell layer. **(I)** Expanded version of the white-framed field of view in **H**. **(J)** Schematic of two-photon calcium imaging of GCs while stimulating MFs with a large-diameter glass electrode. **(K)** Example calcium responses of six GCs recorded using two-photon linescan imaging. The inset shows the imaging field of view and the scanning line drawn over 6 GCs. **(L)** Raster plots of Crus I (left) and Simplex (right) GC responses to 300 Hz MF stimulation. Each row represents the responses of a single GC. White lines indicate stimulus onset and offset. **(M)** Quantification of FWHM of GC responses to 100 and 300 Hz MF stimulation in Crus I (red, 0.93 ± 0.19 s, n = 80 for 100 Hz; 0.98 ± 0.38 s, n = 104 for 300 Hz) and Simplex (blue, 0.96 ± 0.18 s, n = 89 for 100 Hz; 1.17 ± 0.30 s, n = 103 for 300 Hz). Data are presented as median with interquartile range. FWHM is significantly shorter in Crus I at both 100 Hz (Mann-Whitney test, P = 0.03) and 300 Hz (Mann-Whitney test, P < 0.0001) MF stimulation.

Network simulations with dynamic MF-GC synapses reproduce observed GC sparsification and region specificity

To test whether region-specific MF–GC synaptic dynamics are a plausible mechanism for generating sparse, region-specific GC representations, we leveraged a cerebellar cortex firing-

rate model that incorporates experimentally constrained MF–GC STP diversity (*53*) (**Fig. 4A, B**). With minimal parameter tuning, regional differences were captured by selectively increasing the refilling time constant (τ_{ref}) of the low-release-probability vesicle pool (**Fig. 4B**), resulting in Simplex-like synapses having a higher PPR and higher steady-state amplitudes than Crus I-like synapses, without affecting the mean initial EPSC, in qualitative agreement with our experimental findings (**Fig. 2** and **fig. S7C**). This also reproduced experimentally observed differences in synaptic time constants, yielding a 1.25–1.5-fold slowing of τ_{syn} in Simplex (**Fig. 4 C, D,** see **Fig. 2G**). The model was driven with MF activity patterns approximating pontine and Pr5 whisker-air-puff-evoked inputs (**Fig. 1**, see **Materials and Methods**). Simulated transmitter release and GC population activity were markedly sharpened relative to MF input, as quantified by reduced response FWHM (**Fig. 4E, G**). Simplex-like synapses produced broader GC activity profiles than Crus I-like synapses yet remained substantially sharper than their parent MF inputs (**Fig. 4E, G**). In contrast, removing STP resulted in only modest sharpening (**Fig. 4F, G**), suggesting that MF–GC STP is the principal mechanism underlying region-specific and temporally sparse GC activity.

We next asked whether these region-specific GC temporal bases impose distinct constraints on downstream learning. A linear readout unit representing a Purkinje cell (PC; **Fig. 4A**) was trained, using a biologically plausible supervised learning rule, to generate predefined PC output activity profiles from simulated GC activity (see **Materials and Methods**). Incorporating dynamic MF–GC synapses substantially improved learning, enabling both Crus I- and Simplex-like models to reproduce realistic PC target responses (*57*) for target profile onset delays up to ~700 ms (**Fig. 4H**). Strikingly, the two models exhibited complementary learning specializations. Crus I-like synapses preferentially supported learning of brief, early-onset PC targets (**Fig. 4I; fig. S9A-D**), whereas Simplex-like synapses were more effective for broader, late-onset targets (**Fig. 4J**; **fig. S9A-D**). These biases persisted under full gradient descent (**fig. S9A-C**), indicating that they arise from intrinsic differences in GC temporal basis structure rather than properties of the learning rule. In addition, when reducing synaptic diversity, learning performance in broad, late-onset targets was decreased in both Crus I- and Simplex-like models (**fig. S9E-H**). Together, these simulations indicate that MF–GC pre-synaptic dynamics can be regionally tuned to sharpen GC temporal representations and optimize learning to match distinct behavioral time scales.

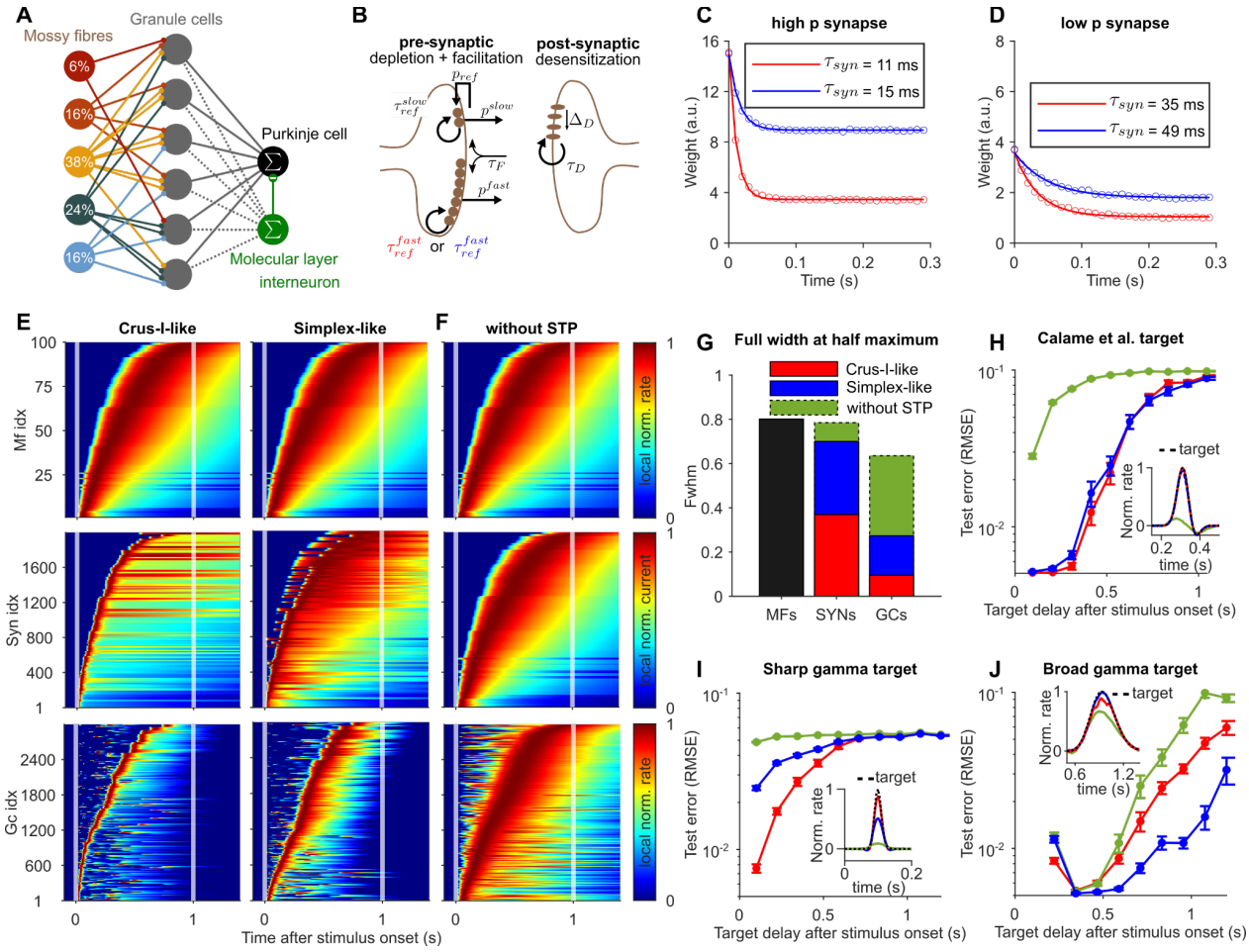

Fig. 4. Model simulations show that region-specific heterogeneous synaptic dynamics can sparsify GC representations differently and introduce learning biases.
(A) Scheme of rate-based cerebellar cortex model including MFs, GCs, one Purkinje cell (PC), and one molecular layer interneuron (MLI), adapted from (*53*). Colors and percentages indicate the relative frequency of MF-GC synapse types; see (*53*). **(B)** Scheme of MF-GC synaptic model showing the principal model parameters, adapted from (*53*). The time constant of the fast-refilling pool is changed to mimic Crus I (red) or Simplex (blue) synapses. **(C)** Simulated average synaptic EPSCs of a high-release-probability synapse in response to a 100 Hz stimulation for Crus I-like (red dots) and Simplex-like synapses (blue dots); solid lines: mono-exponential fits with respective time constants shown in the box. **(D)** Same as **C**, but for a low-release-probability synapse. **(E)** Simulated stimulus-evoked population activity of MFs (top), MF-GC synapses (middle), and GCs (bottom); color-coded MF and GC firing rates as well as MF-GC synaptic currents are individually normalized and sorted according to peak position. Left column: activities for Crus I-like synapses; right column: activities for Simplex-like synapses. **(F)** Same as **E** but for static MF-GC synapses and with an identical GC threshold as in **E**. **(G)** Average FWHM values for populations shown in **E** and **F**. **(H)** Reconstruction error of a brief temporal signal representing forelimb reach kinematics (*57*) at varying target delays for different GC basis sets; each point represents an average over 30 learning simulations. Inset: example of target signal (dashed black line) and simulated PC responses after learning. **(I)** Same as **H**, but for narrow gamma target signals with a SD of 0.01 s. **(J)** Same as **H**, but for broad gamma target signals with an SD of 0.15 s.

373 **Region-specific GC dynamics are inherited from differences in regional neurotransmitter**
374 **release dynamics**

375 If MF–GC STP diversity underlies region-specific GC sequence sparsification, it should be
376 directly reflected in the temporal dynamics of vesicular glutamate release, as shown by our
377 simulations, in which vesicle depletion transforms sustained MF firing into brief, transient synaptic
378 drive. Using a fluorescence-based genetically encoded glutamate sensor (*58*), we directly tested
379 these predictions by monitoring MF glutamate release onto single GC dendritic claws. Expression
380 of the sensor postsynaptically also allowed us to examine the within-cell diversity of
381 neurotransmitter release at all incoming synapses. Because GCs are extremely densely packed,
382 sparse expression is essential to resolve individual MF–GC synaptic connections. We therefore
383 took advantage of a sparse tTA-conditional transgenic line (TCGO-dark), generated by CRISPR-
384 mediated removal of the original YFP tag from the TCGO line (*59*), to express iGluSnFR-A184V
385 in a small subset of GCs in Crus I and Simplex (**Fig. 5A,B**). The decay time of iGluSnFR responses
386 to a single action potential is similar to that of GCaMP8f (20–30 ms (*58*)), indicating that differences
387 in response dynamics cannot be attributed to sensor kinetics. In contrast to GCaMP8f responses,
388 which were a proxy for MF activity, changes in iGluSnFR fluorescence in GC dendritic claws during
389 whisker air puff stimuli reflect short-term alterations in spike-evoked glutamate release from
390 individual MF terminals onto GC dendrites (**Supplementary Movie S1**).

391 Consistent with heterogeneous presynaptic release dynamics, iGluSnFR responses
392 recorded from different claws of the same GC could differ markedly (**Fig. 5A**), mirroring the
393 diversity of synaptic strength and STP observed at MF–GC synapses in brain slice experiments
394 (**Fig. 3**) (*31*). Glutamate responses to repeated presentations of the same whisker stimulus were
395 robust across trials in both Crus I and Simplex (**fig. S10A**). Like GC activity responses,
396 neurotransmitter release response amplitudes were not correlated with whisker air puff responses
397 (**fig. S10**). The time course of the glutamate release was diverse in both Crus I and Simplex and
398 reminiscent of neural sequences in GCs (**Fig. 5B**). The distribution of FWHM of MF-GC synaptic
399 responses was more similar to that of GCs for both Crus I and Simplex than to that of pontine MFs
400 or Pr5 MFs (**Fig. 5C**). Similar to GC neural sequences (**Fig. 1**), glutamate release at MF-GC
401 synapses was significantly briefer in Crus I than in Simplex (**Fig. 5D**), consistent with the finding
402 that synapses in Crus I were faster than those in Simplex. Moreover, glutamate release in Simplex
403 was less temporally sparse than in Crus I, similar to GCs (**fig. S10**). Together, these results
404 demonstrate that sparse, region-specific GC sequences arise from temporally structured
405 neurotransmitter release at MF–GC synapses driven by presynaptic STP diversity.
406

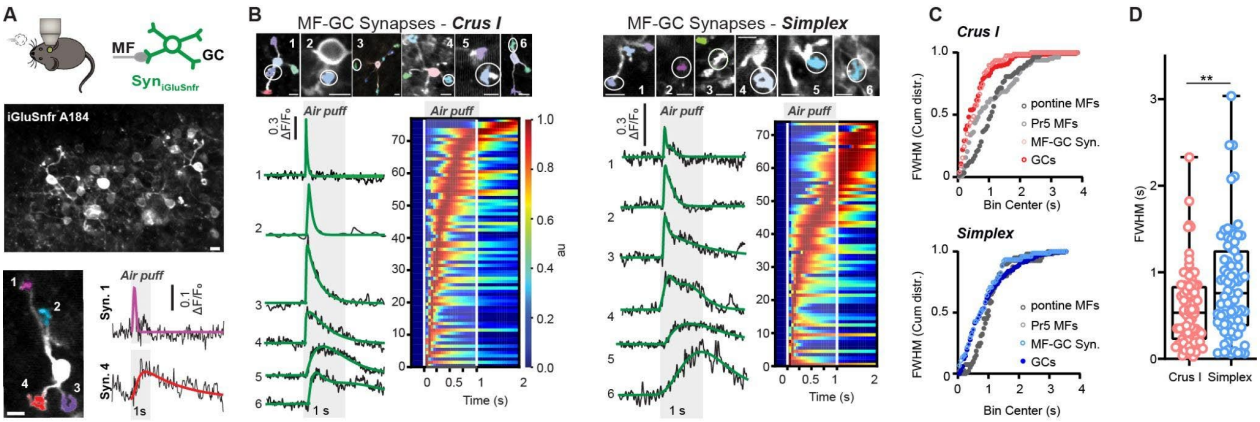

Fig. 5. *In vivo* MF neurotransmitter release dynamics are diverse and faster than MF activity, providing a substrate for precise temporal learning.

(A) Schematic of experimental setup (top). Crus I GCs labeled sparsely with GluSnFR under the TRE promoter in a TCGO mouse (middle). Note that single synapses can easily be identified. Representative example of two MF synapses onto the same GC with different responses (bottom). **(B)** Representative examples of Crus I (left) and Simplex (right) whisker air puff responses from different MF-GC synapses and cells. Note variability across responses. Raster plots of responses of all measured synapses averaged across trials and normalized on the right, respectively (Crus I: $N = 9$, $n = 78$, Simplex: $N = 8$, $n = 75$). **(C)** Cumulative distribution of FWHM for Pr5 MFs, pontine MFs, GCs, and MF-GC synapses. Note the overlap of GC and MF-GC synapse responses, indicating that GCs mimic the temporal response distribution of synaptic responses (two-sample Kolmogorov–Smirnov test, Crus I D = 0.10, $P = 0.66$, Simplex D = 0.13, $P = 0.27$), while it differed between MF-GC synapses and pontine MFs (Crus I D = 0.47, $P < 0.0001$, Simplex D = 0.29, $P = 0.0014$) or Pr5 MFs (Crus I D = 0.86, $P = 0.028$, Simplex D = 0.30, $P = 0.017$). **(D)** The difference in GC whisker air-puff response FWHM between Crus I mimics the difference in MF-GC synapse GluSnFR release between Crus I (0.58 ± 0.05 s) and Simplex (0.85 ± 0.07 s; $P = 0.0049$). Scale bars, 5 µm. Data are shown as median ± interquartile range; values in the text are mean ± SEM.

Discussion

Neural sequences are a ubiquitous coding strategy throughout the brain, yet the circuit and synaptic mechanisms that generate them remain poorly understood. Here, we identify a synaptic mechanism by which the cerebellar input layer transforms sustained sensory drive into temporally sparse population activity that forms a neural sequence tiling the stimulus duration. Cerebellar granule cells (GCs) have long been hypothesized to generate diverse temporal representations of contextual information, which are essential for associative motor learning (*24*, *27*). Using high-speed multiphoton calcium imaging, we showed that air puff stimulation of the whisker field evoked transient GC responses whose peak amplitudes tiled the stimulus, producing a GC activity sequence with narrower temporal profiles than the calcium responses of their mossy fiber (MF) inputs, and that this sharpening is region-specific. Combining *ex vivo* electrophysiology, network modeling, and *in vivo* glutamate imaging, we demonstrated that heterogeneous neurotransmitter release dynamics at MF–GC synapses drive the temporal sharpening of GC activity, are region-specific, and enhance the precision of learning specific time-dependent output activity dynamics. Together, these findings provide direct evidence that heterogeneous synaptic dynamics are a tunable substrate that can sculpt population activity over time to meet the precise timescale demands of cerebellar region-dependent behaviors.

GC neural sequences exhibit multiple computationally relevant properties. Efficient learning systems benefit from internal representations in which complex signals are encoded as combinations of sparsely active units, greatly expanding representational capacity in neural circuits and artificial networks (*60*, *61*). In the cerebellum, sparse GC representations are essential for efficient pattern separation (*18*, *21*, *22*). Our results show that such spatial sparsity is achieved through expansion recoding (*30*, *42*, *43*), whereby fewer than 2% of identified GCs respond to a simple sensory stimulus, compared with a maximum of 16–32% of identified MFs. This spatial sparsity is accompanied by pronounced temporal sparsity within the responsive GC ensemble, providing an additional mechanism for discriminating dynamic patterns of GC activity, as we previously demonstrated for GC delay coding (*31*). The minimal temporal overlap of GC responses yields a population sequence that approximates an orthogonal representation of time, facilitating the learning of time-dependent outputs (see **Fig. 4** and (*7*, *9*, *62*)). The latency of peak GC activity scaled linearly with response width (**Fig. 1**; **fig. S2**), consistent with Weber's scalar timing principle, where variability time estimates increase linearly with the time interval (*63*, *64*). Such Weber-like scaling is a key requirement of theoretical models proposing that cerebellar circuits perform Bayesian estimation of time intervals by learning prior distributions of temporal intervals

(*53*, *62*). Our results, therefore, provide experimental evidence that GC temporal representations satisfy a central prediction of Bayesian models of cerebellar timing. Together, these findings underscore how high–temporal-resolution measurements of neural activity within defined circuit layers can reveal circuit-level computational principles.

Heterogeneous STP sculpts the temporal sparsity of GC neural sequences, thereby expanding the capacity of cerebellar circuits to encode sensory dynamics and support precisely timed motor outputs. Because vesicle replenishment kinetics influence the time course of STP and are subject to neuromodulatory control (*65*, *66*), the dynamic regulation of STP provides a plausible route for adapting the temporal properties of GC activity (*36*) to the behavioral context without altering circuit connectivity. Region-specific synaptic dynamics provide a substrate by which different cerebellar regions could adjust the precision and timescales of temporal learning to be in line with known regional differences in learning rules (*51*). The mechanisms that establish region-specific synaptic timescales could arise through genetic specification, developmental learning, or activity-dependent homeostatic regulation. Regardless of their origin, these differences bias learning in predictable ways: brief GC responses preferentially support learning of fast PC target signals (**Fig. 4H, I**), whereas broader GC responses better support slower target dynamics (**Fig. 4J**). Such biases may align with a mediolateral gradient in cerebellar function, spanning rapid sensorimotor processing to more integrative cognitive operations (*67*). Alternatively, the whisker sensorimotor representations in Crus I (*47*, *68*, *69*) may require fast, temporally sparse representations to support rapid deflections and high spatial acuity, whereas the Simplex region, involved in limb motor control (*70*), may benefit from slower temporal representations to drive motor learning. Notably, the longer GC responses observed in Simplex may also promote generalization by enabling broader, more overlapping temporal representations, potentially accelerating learning across related tasks. Thus, the regional diversity and modulation of MF-GC synaptic dynamics may provide a flexible strategy for the cerebellum to specifically balance temporal precision, bandwidth, and generalization for learning specific behaviors.

Heterogeneity in short-term synaptic plasticity is conserved across animal phyla, from invertebrates to humans (*35*), and implicated in shaping neuronal dynamics and computations (*71*–*74*). Our results position STP diversity within a broader population-level framework by showing that heterogeneous synaptic dynamics provide a biophysical substrate for transforming sensory input into temporally sparse neural sequences tuned for task-specific computations. Given the ubiquitous nature of STP across neural systems, synaptic heterogeneity may be a

fundamental circuit feature that shapes the timing, duration, and distribution of population activity. Such diversity could therefore be tuned to optimize associative learning beyond the cerebellum and influence a wide range of time-dependent computations across sensory, motor, and cognitive domains.

References

1. A. Busch, M. Roussy, R. Luna, M. L. Leavitt, M. H. Mofrad, R. A. Gulli, B. Corrigan, J. Miná̌č, A. J. Sachs, L. Palaniyappan, L. Muller, J. C. Martinez-Trujillo, Neuronal activation sequences in lateral prefrontal cortex encode visuospatial working memory during virtual navigation. *Nat Commun* **15**, 4471 (2024).

2. R. H. R. Hahnloser, A. A. Kozhevnikov, M. S. Fee, An ultra-sparse code underliesthe generation of neural sequences in a songbird. *Nature* **419**, 65–70 (2002).

3. C. D. Harvey, P. Coen, D. W. Tank, Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484**, 62–68 (2012).

4. G. B. M. Mello, S. Soares, J. J. Paton, A Scalable Population Code for Time in the Striatum. *Current Biology* **25**, 1113–1122 (2015).

5. E. Pastalkova, V. Itskov, A. Amarasingham, G. Buzsáki, Internally Generated Cell Assembly Sequences in the Rat Hippocampus. *Science* **321**, 1322–1327 (2008).

6. W. Xie, J. H. Wittig, J. I. Chapeton, M. El-Kalliny, S. N. Jackson, S. K. Inati, K. A. Zaghloul, Neuronal sequences in population bursts encode information in human cortex. *Nature* **635**, 935–942 (2024).

7. S. Zhou, S. C. Masmanidis, D. V. Buonomano, Neural Sequences as an Optimal Dynamical Regime for the Readout of Time. *Neuron* **108**, 651-658.e5 (2020).

8. Q. Yu, Z. Bi, S. Jiang, B. Yan, H. Chen, Y. Wang, Y. Miao, K. Li, Z. Wei, Y. Xie, X. Tan, X. Liu, H. Fu, L. Cui, L. Xing, S. Weng, X. Wang, Y. Yuan, C. Zhou, G. Wang, L. Li, L. Ma, Y. Mao, L. Chen, J. Zhang, Visual cortex encodes timing information in humans and mice. *Neuron* **110**, 4194-4211.e10 (2022).

9. I. R. Fiete, R. H. R. Hahnloser, M. S. Fee, H. S. Seung, Temporal sparseness of the premotor drive is important for rapid learning in a neural network model of birdsong. *J Neurophysiol* **92**, 2274–2282 (2004).

10. S. A. Koay, A. S. Charles, S. Y. Thiberge, C. D. Brody, D. W. Tank, Sequential and efficient neural-population coding of complex task information. *Neuron* **110**, 328-349.e11 (2022).

11. H. Zhu, F. He, P. Zolotavin, S. Patel, A. S. Tolias, L. Luan, C. Xie, Temporal coding carries more stable cortical visual representations than firing rate over time. *Nat Commun* **16**, 7162 (2025).

12. A. P. Vaz, J. H. Wittig, S. K. Inati, K. A. Zaghloul, Replay of cortical spiking sequences during human memory retrieval. *Science* **367**, 1131–1134 (2020).

13. P. R. L. Parker, D. M. Martins, E. S. P. Leonard, N. M. Casey, S. L. Sharp, E. T. T. Abe, M. C. Smear, J. L. Yates, J. F. Mitchell, C. M. Niell, A dynamic sequence of visual processing initiated by gaze shifts. *Nat Neurosci* **26**, 2192–2202 (2023).

14. C. Bimbard, K. D. Harris, M. Carandini, Invariant activity sequences across the mouse brain. bioRxiv [Preprint] (2026). https://doi.org/10.64898/2025.12.20.695676.

15. O. Barak, M. Rigotti, S. Fusi, The Sparseness of Mixed Selectivity Neurons Controls the Generalization–Discrimination Trade-Off. *J. Neurosci.* **33**, 3844–3856 (2013).

16. N. A. Cayco-Gajic, R. A. Silver, Re-evaluating Circuit Mechanisms Underlying Pattern Separation. *Neuron* **101**, 584–602 (2019).

17. M. Rigotti, O. Barak, M. R. Warden, X.-J. Wang, N. D. Daw, E. K. Miller, S. Fusi, The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013).

18. N. A. Cayco-Gajic, C. Clopath, R. A. Silver, Sparse synaptic connectivity is required for decorrelation and pattern separation in feedforward networks. *Nat Commun* **8**, 1116 (2017).

19. W. E. Vinje, J. L. Gallant, Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science* **287**, 1273–1276 (2000).

20. B. A. Olshausen, D. J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**, 607–609 (1996).

21. D. Marr, A theory of cerebellar cortex. *The Journal of Physiology* **202**, 437–470 (1969).

22. J. S. Albus, A theory of cerebellar function. *Mathematical Biosciences* **10**, 25–61 (1971).

23. T. M. Nguyen, L. A. Thomas, J. L. Rhoades, I. Ricchi, X. C. Yuan, A. Sheridan, D. G. C. Hildebrand, J. Funke, W. G. Regehr, W.-C. A. Lee, Structured cerebellar connectivity supports resilient pattern separation. *Nature* **613**, 543–549 (2023).

24. D. V. Buonomano, M. Mauk, Neural network model of the cerebellum: Temporal discrimination and the timing of motor responses. *Neural Computation* **6**, 38–55 (1994).

25. P. Dean, J. Porrill, C.-F. Ekerot, H. Jörntell, The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nature Reviews Neuroscience* **11**, 30–43 (2010).

26. C. Rössert, P. Dean, J. Porrill, At the Edge of Chaos: How Cerebellar Granular Layer Network Dynamics Can Provide the Basis for Temporal Filters. *PLOS Computational Biology* **11**, e1004515 (2015).

27. J. F. Medina, M. D. Mauk, Computer simulation of cerebellar information processing. *Nat Neurosci* **3 Suppl**, 1205–1211 (2000).

28. T.-W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda, D. S. Kim, Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).

29. M. Beau, D. J. Herzfeld, F. Naveros, M. E. Hemelt, F. D'Agostino, M. Oostland, A. Sánchez-López, Y. Y. Chung, M. Maibach, S. Kyranakis, H. N. Stabb, M. G. M. Lopera, A. Lajko, M. Zedler, S. Ohmae, N. J. Hall, B. A. Clark, D. Cohen, S. G. Lisberger, D. Kostadinov, C. Hull, M. Häusser, J. F. Medina, A deep learning strategy to identify cell types across species from high-density extracellular recordings. *Cell* **188**, 2218-2234.e22 (2025).

30. J. I. Gilmer, M. A. Farries, Z. Kilpatrick, I. Delis, J. D. Cohen, A. L. Person, An emergent temporal basis set robustly supports cerebellar time-series learning. *J Neurophysiol* **129**, 159–176 (2023).

31. F. P. Chabrol, A. Arenz, M. T. Wiechert, T. W. Margrie, D. A. DiGregorio, Synaptic diversity enables temporal coding of coincident multisensory inputs in single neurons. *Nat. Neurosci.* **18**, 718–727 (2015).

32. A. Kennedy, G. Wayne, P. Kaifosh, K. Alviña, L. F. Abbott, N. B. Sawtell, A temporal basis for predicting the sensory consequences of motor commands in an electric fish. *Nat Neurosci* **17**, 416–422 (2014).

33. K. Rajan, C. D. Harvey, D. W. Tank, Recurrent Network Models of Sequence Generation and Memory. *Neuron* **90**, 128–142 (2016).

34. R. Laje, D. V. Buonomano, Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nat Neurosci* **16**, 925–933 (2013).

35. J. Beninger, J. Rossbroich, K. Tóth, R. Naud, Functional subtypes of synaptic dynamics in mouse and human. *Cell Reports* **43** (2024).

36. S. Zhou, D. V. Buonomano, Unified control of temporal and spatial scales of sensorimotor behavior through neuromodulation of short-term synaptic plasticity. *Science Advances* **10**, eadk7257 (2024).

37. M. G. Garcia-Garcia, A. Kapoor, O. Akinwale, L. Takemaru, T. H. Kim, C. Paton, A. Litwin-Kumar, M. J. Schnitzer, L. Luo, M. J. Wagner, A cerebellar granule cell-climbing fiber computation to learn to track long time intervals. *Neuron* **112**, 2749-2764.e7 (2024).

38. A. Giovannucci, A. Badura, B. Deverett, F. Najafi, T. D. Pereira, Z. Gao, I. Ozden, A. D. Kloth, E. Pnevmatikakis, L. Paninski, C. I. De Zeeuw, J. F. Medina, S. S.-H. Wang, Cerebellar granule cells acquire a widespread predictive feedback signal during motor learning. *Nat Neurosci* **20**, 727–734 (2017).

39. M. J. Wagner, T. H. Kim, J. Savall, M. J. Schnitzer, L. Luo, Cerebellar granule cells encode the expectation of reward. *Nature* **544**, 96–100 (2017).

40. Y. Zhang, M. Rózsa, Y. Liang, D. Bushey, Z. Wei, J. Zheng, D. Reep, G. J. Broussard, A. Tsang, G. Tsegaye, S. Narayan, C. J. Obara, J.-X. Lim, R. Patel, R. Zhang, M. B. Ahrens, G. C. Turner, S. S.-H. Wang, W. L. Korff, E. R. Schreiter, K. Svoboda, J. P. Hasseman, I. Kolb, L. L. Looger, Fast and sensitive GCaMP calcium indicators for imaging neural populations. *Nature* **615**, 884–891 (2023).

41. J. F. Medina, W. L. Nores, T. Ohyama, M. D. Mauk, Mechanisms of cerebellar learning suggested by eyelid conditioning. *Current Opinion in Neurobiology* **10**, 717–724 (2000).

42. J. S. Albus, A theory of cerebellar function. *Mathematical Biosciences* **10**, 25–61 (1971).

43. D. Marr, A theory of cerebellar cortex. *J. Physiol. (Lond.)* **202**, 437–470 (1969).

44. J. M. Bower, D. H. Beermann, J. M. Gibson, G. M. Shambes, W. Welker, Principles of organization of a cerebro-cerebellar circuit. Micromapping the projections from cerebral (SI) to cerebellar (granule cell layer) tactile areas of rats. *Brain Behav Evol* **18**, 1–18 (1981).

45. L. W. J. Bosman, S. K. E. Koekkoek, J. Shapiro, B. F. M. Rijken, F. Zandstra, B. van der Ende, C. B. Owens, J.-W. Potters, J. R. de Gruijl, T. J. H. Ruigrok, C. I. De Zeeuw, Encoding of whisker input by cerebellar Purkinje cells. *J Physiol* **588**, 3757–3783 (2010).

46. J. U. Henschke, J. M. Pakan, Disynaptic cerebrocerebellar pathways originating from multiple functionally distinct cortical areas. *eLife* **9**, e59148 (2020).

47. R. D. Proville, M. Spolidoro, N. Guyon, G. P. Dugué, F. Selimi, P. Isope, D. Popa, C. Léna, Cerebellum involvement in cortical sensorimotor circuits for the control of voluntary movements. *Nature Neuroscience* **17**, 1233–1239 (2014).

48. B. Willmore, D. J. Tolhurst, Characterizing the sparseness of neural codes. *Network* **12**, 255–270 (2001).

49. J. Morissette, J. M. Bower, Contribution of somatosensory cortex to responses in the rat cerebellar granule cell layer following peripheral tactile stimulation. *Exp Brain Res* **109**, 240–250 (1996).

50. M. Shimuta, I. Sugihara, T. Ishikawa, Multiple signals evoked by unisensory stimulation converge onto cerebellar granule and Purkinje cells in mice. *Commun Biol* **3**, 381 (2020).

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

52. R. E. Thomas, F. Mudlaff, K. Schweers, W. T. Farmer, A. Suvrathan, Heterogeneity in Slow Synaptic Transmission Diversifies Purkinje Cell Timing. *J. Neurosci.* **44** (2024).

53. A. Barri, M. T. Wiechert, M. Jazayeri, D. A. DiGregorio, Synaptic basis of a sub-second representation of time in a neural circuit model. *Nat Commun* **13**, 7902 (2022).

54. P. B. Sargent, C. Saviane, T. A. Nielsen, D. A. DiGregorio, R. A. Silver, Rapid Vesicular Release, Quantal Variability, and Spillover Contribute to the Precision and Reliability of Transmission at a Glomerular Synapse. *J. Neurosci.* **25**, 8173–8187 (2005).

55. E. A. Fleming, G. D. Field, M. R. Tadross, C. Hull, Local synaptic inhibition mediates cerebellar granule cell pattern separation and enables learned sensorimotor associations. *Nat Neurosci* **27**, 689–701 (2024).

56. R. Sosnik, S. Haidarliu, E. Ahissar, Temporal Frequency of Whisker Movement. I. Representations in Brain Stem and Thalamus. *Journal of Neurophysiology* **86**, 339–353 (2001).

57. D. J. Calame, M. I. Becker, A. L. Person, Cerebellar associative learning underlies skilled reach adaptation. *Nat Neurosci* **26**, 1068–1079 (2023).

58. J. S. Marvin, B. Scholl, D. E. Wilson, K. Podgorski, A. Kazemipour, J. A. Müller, S. Schoch, F. J. U. Quiroz, N. Rebola, H. Bao, J. P. Little, A. N. Tkachuk, E. Cai, A. W. Hantman, S. S.-

H. Wang, V. J. DePiero, B. G. Borghuis, E. R. Chapman, D. Dietrich, D. A. DiGregorio, D. Fitzpatrick, L. L. Looger, Stability, affinity, and chromatic variants of the glutamate sensor iGluSnFR. *Nat Methods* **15**, 936–939 (2018).

59. C.-C. Huang, K. Sugino, Y. Shima, C. Guo, S. Bai, B. D. Mensh, S. B. Nelson, A. W. Hantman, Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells. *Elife* **2**, e00400 (2013).

60. D. A. Danhofer, D. D'Ascenzo, R. Dubach, T. Poggio, Position: A Theory of Deep Learning Must Include Compositional Sparsity. arXiv arXiv:2507.02550 [Preprint] (2025). https://doi.org/10.48550/arXiv.2507.02550.

61. T. Poggio, M. Fraser, Compositional sparsity of learnable functions. *Bull. Amer. Math. Soc.* **61**, 438–456 (2024).

62. D. Narain, E. D. Remington, C. I. D. Zeeuw, M. Jazayeri, A cerebellar mechanism for learning prior distributions of time intervals. *Nat Commun* **9**, 469 (2018).

63. J. Gibbon, Scalar expectancy theory and Weber's law in animal timing. *Psychological Review* **84**, 279–325 (1977).

64. J. J. Paton, D. V. Buonomano, The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. *Neuron* **98**, 687–705 (2018).

65. M. R. Carey, W. G. Regehr, Noradrenergic control of associative synaptic plasticity by selective modulation of instructive signals. *Neuron* **62**, 112–122 (2009).

66. E. Fleming, C. Hull, Serotonin regulates dynamics of cerebellar granule cell activity by modulating tonic inhibition. *J Neurophysiol* **121**, 105–114 (2019).

67. C. J. Stoodley, A. M. D'Mello, J. Ellegood, V. Jakkamsetti, P. Liu, M. B. Nebel, J. M. Gibson, E. Kelly, F. Meng, C. A. Cano, J. M. Pascual, S. H. Mostofsky, J. P. Lerch, P. T. Tsai, Altered cerebellar connectivity in autism and cerebellar-mediated rescue of autism-related behaviors in mice. *Nat Neurosci* **20**, 1744–1751 (2017).

68. V. Romano, L. De Propris, L. W. Bosman, P. Warnaar, M. M. ten Brinke, S. Lindeman, C. Ju, A. Velauthapillai, J. K. Spanke, E. Middendorp Guerra, T. M. Hoogland, M. Negrello, E. D'Angelo, C. I. De Zeeuw, Potentiation of cerebellar Purkinje cells facilitates whisker reflex adaptation through increased simple spike activity. *eLife* **7**, e38852 (2018).

69. V. Romano, M. van Driessche, N. van Wingerden, S. Bauer, B. Boeser, J. F. Mejias, C. I. De Zeeuw, Stage-dependent cerebrocerebellar communication during sensorimotor processing. *Nat Commun* **16**, 8812 (2025).

70. K. H. Lee, P. J. Mathews, A. M. B. Reeves, K. Y. Choe, S. A. Jami, R. E. Serrano, T. S. Otis, Circuit Mechanisms Underlying Motor Memory Formation in the Cerebellum. *Neuron* **86**, 529–540 (2015).

71. D. Dahmen, A. Hutt, G. Indiveri, A. Kennedy, J. Lefebvre, L. Mazzucato, A. E. Motter, R. Narayanan, M. Payvand, H. Planert, R. Gast, How heterogeneity shapes dynamics and computation in the brain. *Neuron*, S0896627325008967 (2025).

72. N. Perez-Nieves, V. C. H. Leung, P. L. Dragotti, D. F. M. Goodman, Neural heterogeneity promotes robust learning. *Nat Commun* **12**, 5791 (2021).

73. Y. Wang, H. Markram, P. H. Goodman, T. K. Berger, J. Ma, P. S. Goldman-Rakic, Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat Neurosci* **9**, 534–542 (2006).

74. L. F. Abbott, W. G. Regehr, Synaptic computation. *Nature* **431**, 796–803 (2004).

75. M. i. Aller, A. Jones, D. Merlo, M. Paterlini, A. h. Meyer, U. Amtmann, S. Brickley, H. e. Jolin, A. n. j. McKenzie, H. Monyer, M. Farrant, W. Wisden, Cerebellar granule cell Cre recombinase expression. *genesis* **36**, 97–103 (2003).

76. J. S. Marvin, B. G. Borghuis, L. Tian, J. Cichon, M. T. Harnett, J. Akerboom, A. Gordus, S. L. Renninger, T.-W. Chen, C. I. Bargmann, M. B. Orger, E. R. Schreiter, J. B. Demb, W.-B. Gan, S. A. Hires, L. L. Looger, An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nat Methods* **10**, 162–170 (2013).

77. J. S. Rothman, R. A. Silver, NeuroMatic: An Integrated Open-Source Software Toolkit for Acquisition, Analysis and Simulation of Electrophysiological Data. *Front Neuroinform* **12**, 14 (2018).

78. M. Pachitariu, C. Stringer, M. Dipoppa, S. Schröder, L. F. Rossi, H. Dalgleish, M. Carandini, K. D. Harris, Suite2p: beyond 10,000 neurons with standard two-photon microscopy. bioRxiv [Preprint] (2017). https://doi.org/10.1101/061507.

Funding

Work in the DD laboratory was supported by the Fondation pour la Recherche Médicale (FRM EQU202003010555), Fondation pour l'Audition (FPA-RD-2018-8), the Agence Nationale de la Recherche (ANR-17-CE16-0019, ANR-18-CE16-0018, ANR-19-CE16-0019-02, ANR-21-CE16-0036-01, ANR-23-CE37-0026-03, and ANR-23-CE23-0034-03), and by the NIH (R21 EY026434). FB was additionally supported by the Human Frontier Science Program postdoctoral fellowship (HFSP, LT000674/2019-L), BSS by Pasteur-Roux-Cantarini postdoctoral fellowship from the Institut Pasteur (PRC), SS by the NIH (F32 NS141896), and SBB by a stipend from the Pasteur-Paris University (PPU) International PhD Program. We thank Maureen McFadden for helping with *ex vivo* synaptic time constant analysis. We thank Florian Ruckerl for technical assistance. We thank the Wang lab for providing the AAVDJ-CAG-FLEx-jGCaMP8f construct. We thank the Histology Platform at the Institut Pasteur for providing high-resolution histology images.

Contributions

D.A.D., A.H., F.B., and B.S.S. designed experiments. F.B. performed and analyzed *in vivo* experiments. B.S.S. performed and analyzed *ex vivo* experiments. S.B.B. performed and analyzed *ex vivo* voltage-clamp and part of the cell-attached experiments. A.B. designed and performed model simulations. S.S. performed part of *the in vivo* GC Calcium imaging experiments. M.M.C. performed MLI *in vivo* experiments. G.D. and G.M. conceived the behavioral analysis pipeline. A.H. provided viral constructs for *in vivo* synapse imaging. All authors wrote and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Materials and Methods

Mice

All animal experimental procedures complied with the ARRIVE guidelines and were conducted in compliance with French and European regulations on care and protection of laboratory animals (EC Directive 2010/63, French Law 2013-118, February 6th, 2013) and approved by the Ethics Committee for animal experimentation at the Pasteur Institute (CETEA), or the Institute of Experimental Medicine Protection of Research Subjects Committee, or the University of Colorado Anschutz Medical Campus Institutional Animal Care and Use Committee. Mice were housed in the Pasteur Institute animal facilities, accredited by the French Ministry of Agriculture for performing experiments on live rodents, or at the University of Colorado Anschutz Medical Campus. Mice were kept on a reverse 12 h or 14:10 h light/dark cycle with food and water *ad libitum*. For *in vivo* experiments, GC data came from 14 Gabra6$^{tm2(cre)Wwis}$ (Jackson Laboratory strain 3047798) and 8 B6.D2-Tg(Gabra6-cre)B1Lfr/Mmucd mice, MF data from 9 αGabra6-Cre (B6J-Duke) mice, GC-MF Synapse data from 29 TCGO dark mice (*59*), and MLI data from 7 WT (C57BL/6J, Jackson Laboratory strain 000664) mice. For *in vitro* calcium imaging experiments, data came from 14 alphaGabra6(cre)(B6j-Duke) mice. For *in vitro* synaptic recordings, data were obtained from F1 mice (N = 103, CB6F1 cross of BALB/cJ and C57BL/6J).

Viral injections

For viral injections, mice were first deeply anesthetized with 1 to 4% isoflurane mixed in oxygen. They were then placed in a stereotaxic surgery frame using a mouth clamp. Carprofen (Metacam) was injected intraperitoneally as an analgesic, and mice received additional subcutaneous injections of buprenorphine and dexamethasone. After repeatedly disinfecting the top of the mouse head with liquid betadine and 70% ethanol, lidocaine was injected under the scalp for local anesthesia. A heating blanket with a closed-loop temperature control system was used to maintain body temperature at ~37 °C. Throughout the surgery, temperature and breathing were monitored closely. Eyes were covered with ointment (Viscotears, Alcon, USA; VITA-POS, Pharma Medica AG, Switzerland) to prevent drying. Small scissors were used to open the scalp and expose the skull. Cotton swabs and a scalpel were then used to clean the skull and remove remaining tissue. Lateral muscles on the skull were detached using a scalpel to access the cerebellum. After careful alignment of the skull on the stereotaxic frame, we identified the region of interest for injections. For all targeted structures, coordinates were measured from bregma as

34 follows: Pr5, 5.0 mm posterior, 1.8 mm lateral, 3.5 mm deep from brain surface; BPN, 4.0 mm
35 posterior, 0.4 mm lateral, 5.8, 5.5, 5.2, and 5.0 mm deep from brain surface; Crus I, 6 mm
36 posterior, 3 mm lateral, 0.3-0.7 mm deep from brain surface; Simplex, 6 mm posterior, 1.75 mm
37 lateral, 0.3-0.7 mm deep from brain surface. A small craniotomy of about 0.5 mm diameter was
38 made at the targeted location, and forceps were used to lift the bone cap to access the brain. A
39 thin glass pipette (PCR Micropipets 1-10 µl, Drummond Scientific Company, USA) was first pulled,
40 and then the tip was broken using a tissue to give a 21–27 µm inner tip diameter. The pipette was
41 filled with mineral oil, and then the tip was filled with the AAV vector. The pipette was slowly
42 lowered to the target location in the brain, and injections were performed using an oil-hydraulic
43 injector (Nanoject III, Harvard Apparatus). To specifically express GCaMP8f in GCs, we used the
44 Gabra6-Cre mouse line, which expresses Cre recombinase exclusively in GCs (*75*). We delivered
45 AAVDJ-CAG-FLEx-jGCaMP8f (virus titer: $5x10^{13}$ GC/ml, 1:10 diluted in NaCl), a gift from Wang
46 laboratory (Princeton University), to cerebellar Crus I and/or Simplex at a speed of 2 nL/s, 50 nL
47 per injection site to image GC somas and PFs. In a subset of experiments, we targeted GCs by
48 delivering AAV9-syn-FLEX-jGCaMP8f-WPRE (Addgene, virus titer: $3x10^{13}$ GC/ml, undiluted) in
49 B6.D2-Tg(Gabra6-cre)B1Lfr/Mmucd mice to Crus I and/or Simplex (0.3 and 0.6 mm deep from
50 brain surface), at a speed of 1 nL/s, 150 nL per injection site. To express GCaMP8f in Pr5 and
51 BPN, we used AAV1-hSyn-gCaMP8f ($3x10^{13}$, 1:5 diluted in NaCl). The virus was delivered at 2
52 nL/s, 50 nL per injection site. For MLI experiments, AAV1-syn-GCaMP8f-WPRE (virus titer: 3 x
53 10^{13} GC/ml, 1:10 diluted in NaCl, from the Adam Hantman laboratory, Janelia Research Campus)
54 was delivered at 2 nL/s, 100 nL at 0.2 mm depth. For the sparse expression of SF-iGluSnFR in
55 GCs, we injected AAVDJ.TRE.SF-iGluSnFR.A184V (virus titer: $9.32x10^{12}$ GC/ml, 1:10 diluted in
56 NaCl (*58*, *76*)) in TCGO (*59*) mice that express tetracycline transactivator in a subset of GCs.
57 Injected mice were then kept for 2 weeks to allow transgene expression.
58

59 **Headplate and cranial window surgery for *in vivo* imaging experiments**

60 Animals were given buprenorphine (Buprecare, 0.05-0.1 mg/kg subcutaneously) 30
61 minutes before surgeries. Mice were then deeply anesthetized with isoflurane (4% for induction,
62 1.0–2.5% for maintenance). A 3.5 mm craniotomy was performed above the right hemispheric
63 cerebellum, exposing lobules Crus I and Simplex (6.5 mm anterior to bregma, 2.5 mm lateral to
64 the midline). A window composed of a glass coverslip (Multichannel Systems, CS-3.5R Small
65 round cover glass, #1 thickness) glued (Norland Optical Adhesive 71) to a stainless steel ring
66 (Ziggy's Tubes and Wires; 316 stainless-steel hypodermic tubing, 9R gauge; OD 0.1470–0.1490″,

ID 0.1150–0.1200″; length 0.0197″; cut and deburred) was carefully placed, prefixed with Vetbond (World Precision Instruments), and then glued to the skull with superglue. A one-piece headplate (UCL or Pasteur 3d print facility) was attached around the craniotomy to the animal's head with dental cement (Paladur, Dentalzon). Metacam (1 mg/kg) was given post-surgery, and animals were allowed at least 3 days of recovery before the start of experiments. In a subset of animals, viral injections and cranial window implantation were performed during a single surgery.

Two-photon imaging of head-fixed mice *in vivo*

Mice were head-fixed with custom head bars (UCL) to the permanent head plate and rested in a custom plastic tube (10 cm long, 5 cm inner diameter) to reduce the confounding influence of body movements during experiments. Mice were habituated to the set-up and head fixation in at least 3 sessions of at least 15 minutes each before the start of imaging experiments. Air puffs were produced by a Picospritzer III (Parker Hannifin Corporation) and delivered to the whisker (~1 cm distance, frontal) via a nozzle (1 mm diameter opening) at 5 psi for 1 s, 500 ms or 100 ms. The timing of behavioural images, calcium imaging time series, and whisker air-puff delivery was controlled by triggering TTL pulses with NeuroMatic (*74*) (IGOR Pro, WaveMetrics). Two-photon calcium imaging was performed using a 920 nm laser light generated by a Mai Tai Sapphire laser (Spectra Physics) and a custom-built microscope (2PLSM, Ultima IV, Bruker Nano Surfaces Division, Middleton, WI). The laser beam was scanned across the region of interest through a 16x water-immersion objective (Nikon, 0.8 NA water), which was immersed in ultrasound gel (Sonigel, Mettler Electronics). Excitation power at the back of the objective did not exceed 60 mW. All fluorescence calcium and glutamate transients were imaged in Prairie View using resonant galvo mode. Images were acquired at frequencies between 30 and 232 Hz. Image sizes were 512 x 512 pixels for 30 Hz acquisition rates, and the number of resonant galvo lines was reduced to increase the acquisition rate. A typical experiment consisted of 10 or 30 trials, each trial lasting 15 seconds with stimulation given at 7 seconds (1 s, 0.5 s or 0.1 s whisker air puff). For PF experiments, data were collected within the molecular layer at locations where we observed a high density of labeled PFs. A subset of experiments was performed on a Bruker Ultima 2Pplus using a Vision II laser (Coherent), and a 20x water-immersion objective (Olympus, 1.0 NA), and the timing was controlled using the Prairie View software (Bruker Nano Surfaces Division, Middleton, WI).

Cerebellar slice preparation

Acute coronal slices (200 μm thickness) of cerebellum were prepared from adult Gabra6-

101 Cre mice of either sex, aged between P51-P97, for calcium imaging, and from adult CB6F1 mice
102 of either sex, aged between P30 and P90, for synaptic recordings. Following transcardial
103 perfusion with an ice-cold solution containing (in mM): 2.5 KCl, 0.5 CaCl2, 4 MgCl2, 1.25
104 NaH2PO4, 24 NaHCO3, 25 glucose, 230 sucrose, and 0.5 ascorbic acid, the brains were
105 removed and placed in the same solution. Slices were cut from the dissected cerebellum using a
106 vibratome (Leica VT1200S), and incubated at room temperature for 30 min in a solution
107 containing (in mM): 85 NaCl, 2.5 KCl, 0.5 CaCl2, 4 MgCl2, 1.25 NaH2PO4, 24 NaHCO3, 25
108 glucose, 75 sucrose, and 0.5 ascorbic acid. Slices were then transferred to an external recording
109 solution containing (in mM): 125 NaCl, 2.5 KCl, 1.5 CaCl2, 1.5 MgCl2, 1.25 NaH2PO4, 24
110 NaHCO3, 25 glucose, and 0.5 ascorbic acid, and maintained at room temperature for up to 6 hr.
111 All solutions were bubbled with 95% O2 and 5% CO2.
112

113 **Two-photon imaging *in vitro***

114 The brain slices containing GCaMP8f-expressing GC somata and MF terminals were
115 identified with a 4x objective lens (Olympus UplanFI 4x, 0.13 NA) using very brief illumination with
116 470 nm light to excite GCaMP8f fluorescence. GCs were identified using infrared Dodt-gradient
117 contrast and a QIClick digital CCD camera (QImaging, Surrey, BC, Canada) mounted on an
118 Ultima multiphoton microscopy system (Bruker Nano Surfaces Division, Middleton, WI, USA) that
119 was mounted on an Olympus BX61W1 microscope, equipped with a water-immersion objective
120 (Olympus 60x, 1.1 NA). Two-photon excitation was performed with a Ti-sapphire laser
121 (Spectraphysics). To visualize GCs expressing GCaMP8f, two-photon excitation was performed
122 at 920 nm. Infrared Dodt-gradient contrast was used to position the stimulation pipette in the white
123 matter, targeting MFs to indirectly activate GC bodies. Linescan imaging of GC somata and MF
124 terminals was performed by scanning through the cell membrane marked by freehand linescan
125 mode (Prairie View). Total laser illumination per sweep lasted 2000 ms. Fluorescence was
126 detected using both proximal epifluorescence and a substage photomultiplier tube gallium
127 arsenide phosphide (H7422PA-40 SEL, Hamamatsu).
128

129 **Synaptic recordings**

130 Whole-cell patch-clamp recordings of GCs were performed using a Multiclamp 700B
131 amplifier (Axon Instruments) and according to the methodology described in Chabrol et al. (*31*).
132 Thick-walled, fire-polished glass patch electrodes (1.5 mm OD, 0.75 mm ID, Sutter Instruments;
133 5–8 MΩ tip resistance) were filled with a K^+-based internal solution containing (in mM): 110 KOH,
134 110 methanesulfonic acid, 6 NaOH, 4.6 $MgCl_2$, 5 EGTA, 1.78 $CaCl_2$, 4 Na-ATP, 0.3 Na-GTP (280-

285 mOsm, pH 7.3). The Crus I or Simplex region was first identified under a 4x objective lens (Olympus UplanFl 4x, 0.13 NA). GCs were then visualized using infrared Dodt-gradient contrast on a multiphoton microscopy system as described above. Cells were voltage-clamped at -80 mV to record EPSCs elicited by electrical stimulation of single MF inputs. MF transmission was evoked using a second patch pipette filled with the recording solution and randomly placed within the GC layer. Single inputs were isolated with 10 µs voltage pulses (Digitimer Ltd, Letchworth Garden City, UK), and stimulation intensity was set to 5 V above the threshold where consistent synaptic responses could be evoked following each pulse in a 100 Hz train. The AMPA-receptor-mediated component of the synaptic current was isolated using SR 95531 (10 µM) to block $GABA_A$Rs, d-AP5 (10 µM), and 7-chlorokynurenic acid (20 µM) to block NMDA receptors (NMDARs), and strychnine (0.3 µM) to block glycine receptors. Membrane capacitance and series resistance were estimated online by manual cancellation of the transient current responses to a 5 mV step depolarization. Recordings were performed at 35-37°C, low-pass filtered at 10 kHz, digitized at 100 kHz using an analog-to-digital converter (model NI USB 6259, National Instruments, Austin, TX, USA), and acquired using the Clamp module on NeuroMatic (77), running on Igor PRO (Wavemetrics, Lake Oswego, OR, USA).

EPSCs analysis and estimation of synaptic time constants

Electrophysiological data were analyzed using custom scripts in NeuroMatic (Igor Pro 8). Extracellular stimulation artifacts were subtracted using the Art module, with a single-exponential fit applied to the decay, and baseline correction enabled to account for residual EPSC activity. Series resistance and membrane capacitance were corrected using the online estimates, and traces were low-pass filtered at 4 kHz. Input resistance, estimated from the peak of a small capacitive transient preceding each sweep, was used to monitor cell health. Trials failing the NeuroMatic first-pass stability test for input resistance or for the amplitude of the first EPSC in the train (used as a plasticity proxy) were excluded. Average EPSC amplitude was measured within a 100 µs window centered on the mean peak response. CV was calculated as the square root of baseline-subtracted variance divided by the mean amplitude. PPR was computed as EPSC2/EPSC1 from average 100 Hz trains.

Synaptic decay time constants were estimated by fitting single- and double-exponential models to average EPSC decay waveforms recorded during 20, 100, and 300 Hz stimulation trains. Train durations (3,000 ms at 20 Hz; 500 ms at 100 Hz; 250 ms at 300 Hz) were selected to ensure steady-state responses (yielding 61, 51, and 84 stimuli, respectively). EPSC peaks were detected as absolute minima following each pulse. A custom iterative fitting algorithm

169 selected optimal parameters for each model. Fits were evaluated based on: (1) chi-square values;
170 (2) amplitude ratio of the two components (<10-fold difference); and (3) time constant separation
171 (>2-fold difference). If criteria were met, a weighted decay time constant was calculated from the
172 double-exponential fit as $\frac{A_1\tau_1+A_2\tau_2}{A_1+A_2}$. Otherwise, the time constant of a single-exponential model
173 was used. Fits failing to converge—e.g., due to linear EPSC decay—were excluded. Linearity
174 was assessed using the NeuroMatic first-pass stability test and confirmed in GraphPad via slope
175 significance testing.
176

177 **Cell-attached recordings and spike analysis**

178 A large-diameter stimulation electrode (>20 μm) was positioned in the white matter to
179 activate MF tracts. Spiking responses were recorded in the cell-attached configuration during 1s
180 trains of 100 Hz stimulation. Pulse width was initially set to 100 μs and adjusted for each cell to
181 optimize activation. Stimulation intensity was increased to maximize the number of MFs recruited,
182 thereby maximizing evoked spike output. When stable access permitted, whole-cell configuration
183 was subsequently established to record the underlying synaptic drive (EPSCs) under identical
184 stimulation conditions. Stimulation artifacts were blanked offline, and spikes were detected using
185 the Spikes module in NeuroMatic to generate peristimulus time histograms (PSTHs). PSTHs were
186 fitted with single- or double-exponential functions to extract spike decay time constants,
187 analogous to what was described for synaptic time constants.
188

189 **Data preprocessing**

190 Image registration, Neuropil extraction, movement correction, and cell detection of *in vivo*
191 two-photon imaging data were performed in Suite2p (*78*). For GC and MF recordings, cells were
192 divided into active and inactive groups during the recording session, and only active cells—
193 identified manually based on morphology and activity traces—were further analyzed. Trials were
194 curated for movement artifacts: y-shifts of recording at each timepoint, y-off, as determined by
195 Suite2p, had to be below factor 20. Only cells with a maximum value of the smoothed trace
196 (Savitzky-Golay filter, window length = 9, polynomial order = 2) that exceeded 3 standard
197 deviations of the baseline fluorescence (defined as 1 s before stimulation) during stimulus
198 presentation were considered as responding to the stimulus. Each GC dendritic claw was
199 assigned a unique ID based on morphology and imaging region.
200

201 **Imaging data analysis**

202 Data was stored in a MySQL database for further analysis with custom-written Python and

MATLAB scripts. Delta F/F_0 was calculated using a baseline fluorescence period (1 s) immediately prior to stimulation. Calcium traces, 5 seconds long from the start of the stimulus, were fitted by a multi-exponential function with time delay t_0 one rising (τ_R) and two decaying components (τ_{D1}, τ_{D2}) with weights A_1 and A_2, respectively (IGOR Pro, WaveMetrics):

$$y(t) = \left(1 - e^{-\frac{t-t_0}{\tau_R}}\right)^n \cdot \left(A_1 e^{-\frac{t-t_0}{\tau_{D1}}} + A_2 e^{-\frac{t-t_0}{\tau_{D2}}}\right).$$

Time to peak, rise time (to 90% of the peak), decay time (time from peak to 50% of maximum), onset time (to 10% above baseline), and FWHM (time from 50% rise to 50% decay) were determined for each cell from their multiexponential fit.

Temporal sparsity was defined as in (*48*):

$$s = \frac{1 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2 \left(\sum_{i=1}^{n} y_i^2\right)^{-1}}{1 - \frac{1}{n}},$$

where n is the number of recording time steps.

For *ex vivo* calcium imaging, initial image processing and registration were automated using a custom Python-based pipeline. Individual regions of interest (ROIs) corresponding to GC somata were subsequently extracted from the linescan data using a custom-built viewer in RStudio. During data acquisition, freehand scanning lines were positioned across groups of 5–10 labeled cells per sweep to capture simultaneous activity patterns. To differentiate calcium transients from baseline noise, an automated inclusion criterion was implemented where a cell was considered responsive only if its maximum deflection during the stimulus exceeded 3 standard deviations of the baseline fluorescence. From these traces, we determined the FWHM as described above.

To characterize the specific calcium dynamics (kernels) of individual GC somas, MF terminals, and PF boutons, action potential-evoked transients were recorded in response to 1, 2, and 10 pulses at 100 Hz *ex vivo*. Analysis of these dynamics was performed using Igor Pro and NeuroMatic, where twenty trials were collected for each stimulation condition alongside non-stimulated baseline traces used for subtraction and division of the stimulus trials to generate bleach-corrected $\Delta F/F_0$ profiles. Kinetic parameters, including peak amplitude, rise time, and 50% decay time, were directly determined from these corrected traces to compare the inherent calcium signaling properties across neuronal compartments.

Behavioural analysis

Behavioural movies were acquired using two Basler Cameras (ace acA1440-220um USB3 Monochrome, 1440 px x 1080 px resolution), one was filming the mouse head and whiskers at

250 Hz with 500 µs exposure time, the second was filming body movements at 100 Hz. A Python-based custom-written camera driver was used to select optimized image parameters and record single-frame timestamps with microsecond precision. Illumination was provided by an infrared lamp (CMVision IR30 WideAngle IR Illuminator) at 850 nm. ROIs were defined from the video frame to reach the desired recording frequency. Each frame had a unique ID and timestamps up to microsecond precision were automatically recorded. In a subset of experiments, whisker movements were captured at 150 Hz with 500 µs exposure time using a FLIR Blackfly camera (Edmund Optics). To extract whisker kinematics, we used a custom-written Python-based whisker detection and tracking software. The angles of deflection of the left and right whisker arrays were calculated for each image frame within a manually selected radial region around the centre of the whisker pad. Trial-averaged whisker responses to air puff were fitted as for the Ca^{2+} traces using the multi-exponential function described above.

Histology

Following completion of the behavioral experiments, mice were deeply anesthetized with a mixture of Ketamine (Imalgen, 100 mg/kg, Merial) and Xylazine (Rompun, 10 mg/kg, Bayer) via intraperitoneal injection, and transcardially perfused with PBS, followed by 4% paraformaldehyde (PFA) in PBS. Brains were removed and post-fixed overnight in 4% PFA, then transferred to PBS. Following post-fixation, 50 µm whole-brain sagittal sections were made using a vibratome (Leica Biosystems, V1000S). The sections were then mounted in DAPI (1:5000 dilution) Fluoromount (F4680; Sigma-Aldrich), and the slides were sealed at the margins 24h later with nail varnish and stored at 4° Celsius. Fluorescence images of whole brain slices were captured using a laser-scanning confocal microscope with a 488 nm laser for GFP excitation and a 555 nm laser for tdTomato excitation at 20X magnification of 500 μm^2 tiles (Axio Scan.Z1, Inverted LSM 700, Zeiss; from the Histology Platform at Institut Pasteur). The images of the brain slices were then matched to the closest brain cartography from the Allen Brain Atlas.

Statistical analysis

Data plots were generated, and statistical comparisons were performed using the appropriate test in GraphPad Prism version 10.4.1 for Windows (GraphPad Software, Boston, Massachusetts, USA, www.graphpad.com). *P*-values < 0.05 were considered to indicate significance. In summary plots, data are represented as median with interquartile ranges, and box plot whiskers show the full range of values. In figure legends, values indicate median ± standard deviation, except where otherwise noted.

Cerebellar cortex model and simulations

Simulations were carried out using a modified version of the cerebellar cortex model developed in (*53*), to which we refer the reader for a detailed description. Below, we give a brief overview and highlight the differences between the original model and the one used here.

MF-GC synapses

Model synapses were identical to Barri et al. (*53*), except for the fast refilling time constant (see below). The average synaptic current transmitted per unit time by a single MF to a single GC is:

$$I_{syn}(t) = W(t) \cdot m(t),$$

where W(t) is the time-dependent MF-GC synaptic weight and m(t) a continuous variable denoting the instantaneous MF firing rate. W(t) depends on pre-synaptic mechanisms, determining the average number of released vesicles in response to m(t), and a term for post-synaptic receptor desensitization (**Fig. 4B**). Each MF-GC synapse featured two vesicle pools, which differ in their refilling rates and which we refer to as 'slow' and 'fast', respectively. Thus:

$$W(t) = q(t) \cdot \left(n^{slow}(t) + n^{fast}(t)\right),$$

where n^{slow} and n^{fast} are the average number of vesicles released by the two pools, respectively, and q(t) is a post-synaptic factor.

The slow pool was composed of a comparatively small number of docked vesicles, N^{slow}, with a high vesicle release probability $p^{slow,}$ and a low rate of recovery from vesicle depletion $(\tau_{ref}^{slow})^{-1}$. In addition, the slow pool's recovery from depletion was augmented by a process of activity-dependent fast vesicle refilling, captured by the probability p_{ref}. The fast pool had a comparatively large number of docked vesicles, N^{fast}, with a low release probability p^{fast} and a high rate of recovery from depletion $(\tau_{ref}^{fast})^{-1}$. We thus have: $\tau_{ref}^{slow} \gg \tau_{ref}^{fast}$, $N^{slow} < N^{fast}$ and $p^{slow} > p^{fast}$.

For a single synapse, the fraction of neurotransmitter available at the slow and fast pools, x^{slow} and x^{fast} is given by:

$$\dot{x}^{slow}(t) = \frac{1 - x^{slow}(t)}{\tau_{ref}^{slow}} - u^{slow}(t) \cdot (1 - p_{ref}) \cdot x^{slow}(t) \cdot m(t)$$

$$\dot{x}^{fast}(t) = \frac{1 - x^{fast}(t)}{\tau_{ref}^{fast}} - u^{fast}(t) \cdot x^{fast}(t) \cdot m(t).$$

The variables $u^{slow}(t)$ and $u^{fast}(t)$ denote the pools' respective release probabilities, which are time-dependent and modified by short-term facilitation:

$$\dot{u}^{slow}(t) = \frac{p^{slow} - u^{low}(t)}{\tau_F} + p^{slow} \cdot (1 - u^{slow}(t)) \cdot m(t)$$

$$\dot{u}^{fast}(t) = \frac{p^{fast} - u^{fast}(t)}{\tau_F} + p^{fast} \cdot (1 - u^{fast}(t)) \cdot m(t),$$

with τ_F denoting the facilitation time-constant. With this, the average number of vesicles released by the two pools can be written as:

$$n^{slow}(t) = N^{slow} \cdot u^{slow}(t) \cdot x^{slow}(t)$$

$$n^{fast}(t) = N^{fast} \cdot u^{fast}(t) \cdot x^{fast}(t).$$

Finally, post-synaptic receptor desensitization depends on the number of released vesicles according to:

$$\dot{q}(t) = \frac{1 - q(t)}{\tau_D} - \Delta_D \cdot q(t) \cdot \frac{n^{slow}(t) + n^{fast}(t)}{N^{slow} + N^{fast}} \cdot m(t).$$

In the absence of STP (**Fig. 4F**), we set $x^{slow}(t) = x^{fast}(t) = q(t) = 1$, $u^{slow}(t) = p^{slow}$ and $u^{fast}(t) = p^{fast}$ for all t.

Synaptic parameters

As in Barri et al. (*53*), synaptic parameters were constrained to reproduce the average behaviour of the five MF-GC synapse types that were experimentally determined (*31*). All synaptic parameters, except τ_{ref}^{fast}, were unchanged from the original model. To obtain region-specific models, we set the fast pool of each synapse to have a $\tau_{ref}^{fast} = 40$ ms for Crus I-like simulations, and $\tau_{ref}^{fast} = 10$ ms for Simplex-like simulations. This resulted in Simplex-like synapses having a higher PPR, longer and higher steady-state amplitudes than Crus I-like synapses, without affecting the initial mean EPCS, in qualitative agreement with our experimental findings (**fig. S7C**). Note that alternatively lowering/increasing release probabilities to obtain Simplex-like/Crus I-like synapses would yield similar qualitative differences in PPR, τ_{syn} and steady-state amplitudes. However, this would also lead to different initial mean EPSC, which is inconsistent with the experimental data.

Mossy Fibers

In our original model, MFs exhibited constant firing rates, dictated by the group membership of their synapses, that experienced an instantaneous frequency switch at stimulus

onset (*53*). In contrast, here we modelled air-puff stimulus-evoked activity as multi-exponential functions, with one rising (τ_0) and two decaying components (τ_1, τ_2) to approximate the unobserved firing rates underlying our measured calcium traces. Furthermore, the experimentally observed differences between pontine and Pr5 MFs were qualitatively captured by using two families of firing-rate time profiles.

The rate of the ith (out of M) MF was modelled as:

$$m_i^\alpha(t) = 0 \qquad \text{if } t < \Delta t_i^\alpha$$

$$m_i^\alpha(t) = m_i^{max} \cdot \left(1 - e^{-\frac{t-\Delta t_i^\alpha}{\tau_{0,i}^\alpha}}\right) \cdot \left(\gamma_i \cdot e^{-\frac{t-\Delta t_i^\alpha}{\tau_1^\alpha}} + (1-\gamma_i) \cdot e^{-\frac{t-\Delta t_i^\alpha}{\tau_2^\alpha}}\right) \qquad \text{else,}$$

where m^{max} is the peak firing rate, Δt the onset time delay and $0 < \gamma < 1$. The index $\alpha \in \{\text{pons}, \text{Pr5}\}$ indicates pontine or Pr5 origin. MF inputs were associated to the synapse types defined in (*31*) according to the following rule: all synapse types, regardless of their release probability, received pontine MF inputs. Synapses with the highest release probabilities (groups 1 and 2 in (*31*), constituting 22% of all synapses), received Pr5 inputs with probability 0.5, as primary afferents preferentially form high p synapses (*31*).

In keeping with the high firing rates reported in response to whisker stimulation (*56*), peak rates m^{max} were drawn individually for each MF and irrespective of synapse group, and for both pons and Pr5, according to:

$$m_i^{max} \sim \mathcal{N}(200\text{Hz}, (20\text{Hz})^2).$$

MF onset delays were drawn from exponential distributions

$$\Delta t_i^\alpha \sim Exp(\lambda^\alpha),$$

with rate parameters $\lambda^{pons} = 200$ ms for pontine MFs and $\lambda^{Pr5} = 10$ms for Pr5 MFs.

Decay time-constants were set to $\tau_1 = 100$ms, $\tau_2^{pons} = 1000$ms and $\tau_2^{Pr5} = 500$ms to capture the longer duration of pontine MF activation profiles. To account for the temporal variability between MFs, rise time-constants were assigned to be uniformly spaced across M MFs according to $\{\tau_{i,pons}^0\} \in [50, ..., 200]$ms and $\{\tau_{i,Pr5}^0\} \in [10, ..., 100]$ms. Similarly, we set $\{\gamma_i\} \in [0.3, ..., 0.9]$. To capture the experimental observation that that time to peak was positively correlated with the temporal duration of MF activity (FWHM; **fig. S2**), we introduced rank correlations of 1 between $\{\Delta t_i^\alpha\}$ and $\{\tau_{i,\alpha}^0\}$ as well as $\{\Delta t_i^\alpha\}$ and $\{\gamma_i\}$.

358 **Cerebellar cortical circuit model**

359 Our model consisted of firing rate units corresponding to M=100 MFs, N=3000 GCs, a
360 single molecular layer interneuron (MLI) and a single PC that linearly summed excitatory inputs
361 from GCs and MLI inhibition. Each GC received four MF synapses randomly selected according
362 to their relative abundance (*31*). The ith GC input $h_i(t)$ and GC firing rate $g_i(t)$ were determined by:

363 $$\tau_g \, \dot{h}_i(t) = -h_i(t) + \sum_{j \in K} W_{ij}(t) m_j(t)$$

364 $$g_i(t) = [h_i(t) - \theta_i]_+,$$

365 where $[\cdot]_+$ the linear rectifier, $\tau_g = 10$ms, and K is a set of four MF indices, randomly drawn as
366 explained in (*53*). GC thresholds θ_i were adjusted such that each GC was active for a fraction
367 f_c=0.5 of all time steps during the interval from stimulus onset to T_{post}=1.4 s post-onset. In the
368 absence of STP (**Fig. 4F**), GC thresholds were set identically to those obtained in the presence
369 of STP.

370 To compute the PC firing rate, each GC firing-rate profile was independently normalized
371 to a peak value of one. MLI activity was assumed to represent the average rate of the GC
372 population, yielding the PC firing rate:

373 $$pc(t) = \frac{1}{N} \sum_{i=1}^{N} (J_i - 1) \bar{g}_i(t) \quad ,$$

374 where $\bar{g}_i$ are the normalized GC firing rates and J_i are the non-negative GC-PC synaptic weights.
375 Normalizing GC firing rates further simplifies our original model.
376 Model equations were numerically integrated using the Euler method with step size 0.5 ms.
377

378 **Target signals**

379 Target signals used for PC training were either a realistic PC firing rate profile from (*57*)
380 (**Fig. 4H**) or firing profiles modeled by a gamma distribution according to:

381 $$f(t) \propto t^{k_\gamma - 1} e^{-\frac{t}{\theta_\gamma}},$$

382 with normalization chosen such that max(f)=1. In **Fig. 4I and 4J** we adjusted k_γ and θ_γ such that
383 the averages of the resulting Gamma shapes corresponded to the target delay and its SDs to
384 0.01 s and 0.15 s, respectively.
385

386 **Learning rule**

387 Reconstruction of the target signal f(t) was achieved by adjusting the weights J_i via a
388 supervised learning rule. The least squares loss to be minimised by the learning procedure is:

$$E = \frac{1}{2}\int_{-T_{pre}}^{T_{post}} dt\,(pc(t_k) - f(t_k))^2,$$

where T_{pre}= 0.1 s and T_{post}= 1.4 s. Learning via gradient descent leads to a stepwise weight update according to:

$$J_i \leftarrow J_i + \Delta J_i$$

$$\Delta J_i = \eta \frac{\partial \epsilon}{\partial J_i}$$

$$= \frac{\eta}{N} \cdot \int_{-T_{pre}}^{T_{post}} dt\,(pc(t_k) - f(t_k)) \cdot \bar{g}_i(t),$$

where $\eta = 0.0025$ is a learning rate. To increase biological realism, we modified the above by relaying the error information through a non-negative and bounded climbing fibre rate:

$$cf(t) = cf_{max} \tanh\left(\frac{[cf_{spont} + \beta(pc(t) - f(t))]_+}{cf_{max}}\right),$$

leading to the modified weight update:

$$\Delta J_i = \frac{\eta}{N} \cdot \int_{-T_{pre}}^{T_{post}} dt(cf_{spont} - cf(t)) \cdot \bar{g}_i(t).$$

Here, $cf_{spont} = 1$ Hz is the spontaneous climbing fibre rate. Relative to the original implementation, we introduced a maximum climbing fibre rate of $cf_{max} = 20$ Hz and changed the proportionality factor β from 0.5 to 200 to account for the different target signals used here.

As in (*53*), the full gradient update rule (**fig. S9A-C**) and the biologically inspired update rule (**Fig. 4** and **fig. 9D-H**) were modified by a Nesterov acceleration scheme to increase learning speed. For weight learning, we subsampled the simulated and normalized GC rates by a factor of 10 and carried out 4000 update steps. The initial weight distribution was set to J_i=0 for all i. To calculate the reconstruction error, we computed the final PC output by adding zero-mean white noise with a standard deviation of 0.01 to the normalized GC firing rates.

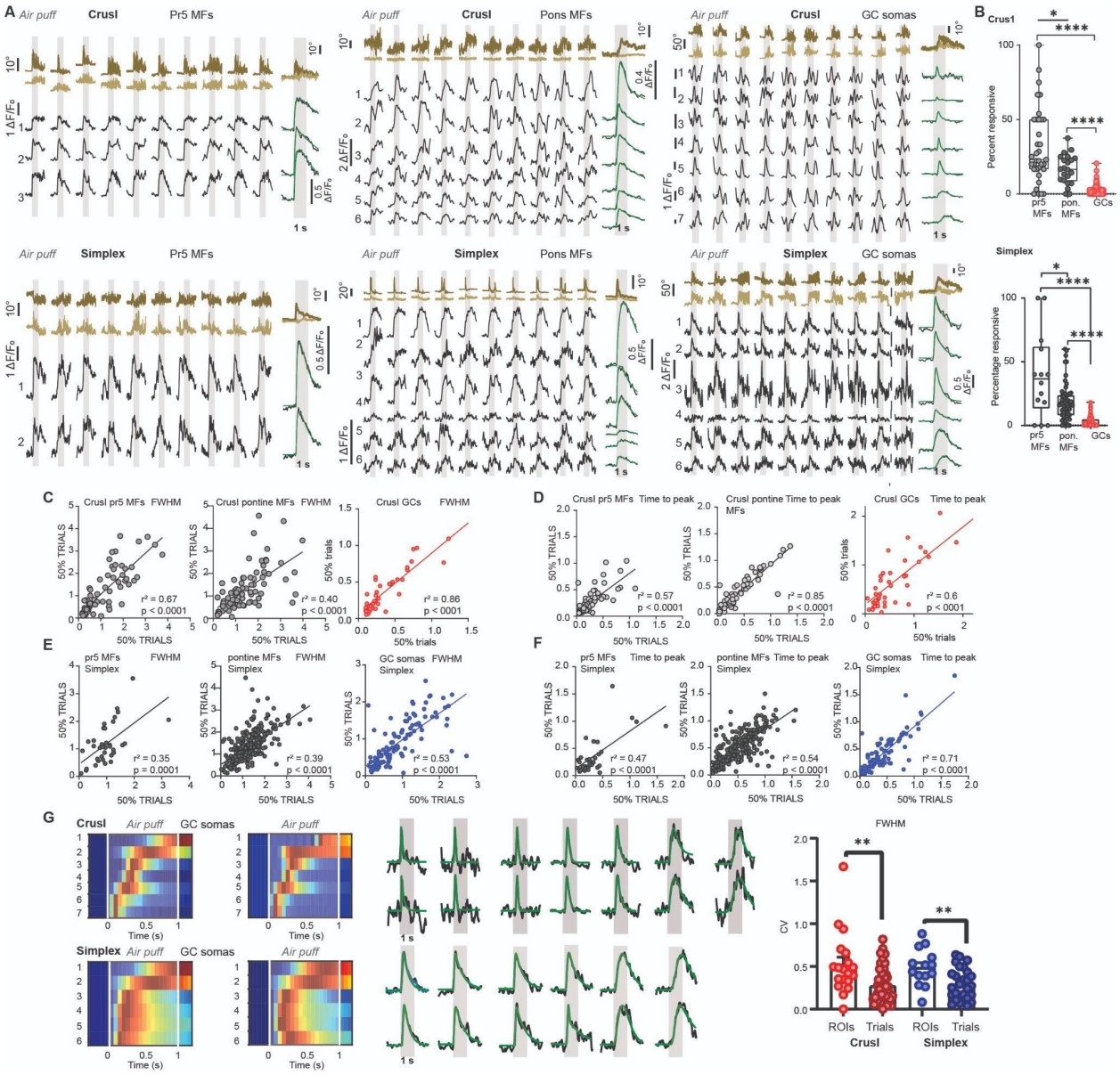

Fig. S1. Region-specific response robustness of Pr5 MF, pontine MF, and GC populations. (A) Calcium responses in Pr5 mossy fibres (MFs), pontine MFs and granule cell (GC) somata in Crus I and Simplex during ten 1-s whisker stimulation trials. **(B)** Fraction of cells responding to whisker air puff relative to all segmented cells recorded during each experiment. In Crus I, 32 ± 4% of Pr5 mossy fibres (MFs; *n* = 35 experiments, *N* = 4 mice), 16 ± 2% of pontine MFs (*n* = 22, *N* = 3), and 2 ± 0.3% of granule cell (GC) somata (*n* = 132, *N* = 6) responded to whisker air puff. In Simplex, 39 ± 9% of Pr5 MFs (*n* = 14, *N* = 3), 18 ± 2% of pontine MFs (*n* = 58, *N* = 3), and 3 ± 0.4% of GC somata (*n* = 83, *N* = 7) were responsive. **(C-F)**, Correlation of full width at half maximum (FWHM) and time to peak calculated from randomly selected and averaged 50% of trials within each experiment in Crus I (**C,D**) and Simplex (**E,F**). **(G)** Averaged, normalized calcium responses obtained by randomly averaging 50% of trials, respectively, of one representative experiment. Right, the coefficient of variation (CV) is lower within the same cell across trials than across different cells within the same experiment (Simplex, *P* = 0.007; Crus I, *P* = 0.001). Experiments were included in the analysis in which at least 3 cells showed whisker-air-puff responses. Data are presented as median ± interquartile range in the figures. Data are reported in text as mean ± SEM.

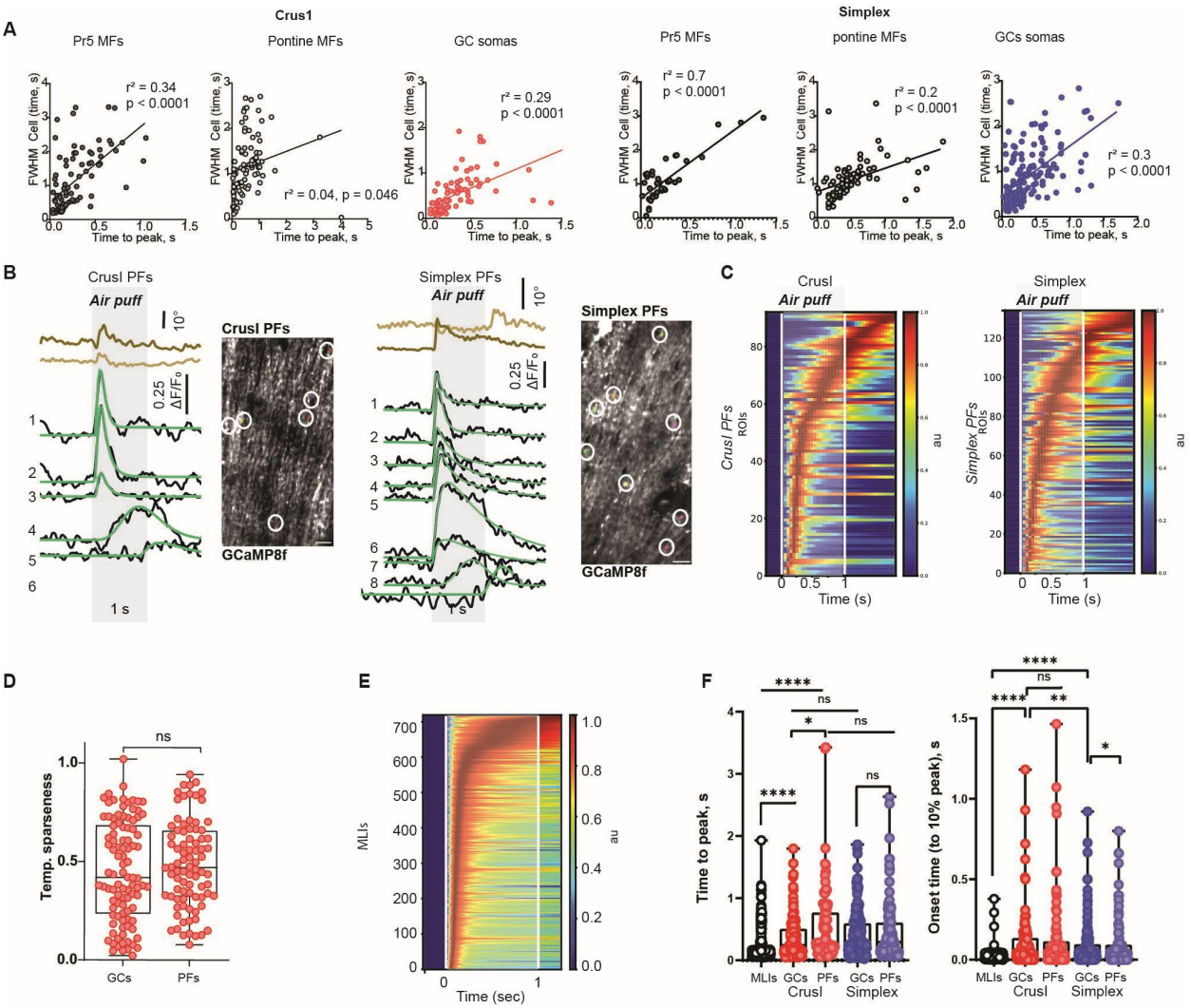

Fig. S2. Temporal dynamics of whisker-air-puff-evoked calcium responses across cerebellar inputs, GC somas and PFs, and MLI interneurons

(A) FWHM correlated with time to peak in Crus I Pr5 MFs (n = 90, N = 4), pontine MFs (n = 89, N = 3), and GCs (n = 90, N = 5). In Simplex, FWHM also correlated with time to peak in pontine MFs (n = 88, N = 3), GCs (n = 156, N = 4), and Pr5 MFs (n = 40, N = 4). **(B)** Averaged calcium responses of six Crus I granule cell PFs and eight Simplex PFs during ten whisker stimulation trials. **(C)** Raster plots showing PF responses in Crus I (n = 92 cells, N = 5 mice) and Simplex (n = 134, N = 6). **(D)** Temporal sparseness did not differ between Crus I GC somata and GC PFs (P = 0.316). **e**, Rasterplot of Crus I molecular layer interneuron (MLI) calcium responses to 1-s whisker stimulation (n = 720, N = 7). **(F)** Time to peak in Crus I MLIs (0.20 ± 0.01 s) was earlier compared with Crus I GC somata (0.36 ± 0.03 s, P < 0.0001) and PFs (0.53 ± 0.06 s, P < 0.0001). Time to peak did not differ between Crus I and Simplex GCs (0.4 ± 0.03 s, P = 0.13) nor between Simplex GCs and PFs (0.45 ± 0.04 s, P = 0.92). The time to peak was earlier in Crus I GC somas than in PFs (P = 0.015). Onset times were earlier in Crus I MLIs (0.02 ± 0.001 s) compared to Crus I GC somata (0.11 ± 0.01 s, P < 0.0001), did not differ between Crus I GC somas and PFs (0.14 ± 0.03 s, P = 0.43), were earlier in Simplex GCs (0.08 ± 0.01 s) than Crus I GCs (P = 0.01) and did not differ between Simplex PFs (0.09 ± 0.01 s) and Simplex GC somas (P = 0.83).

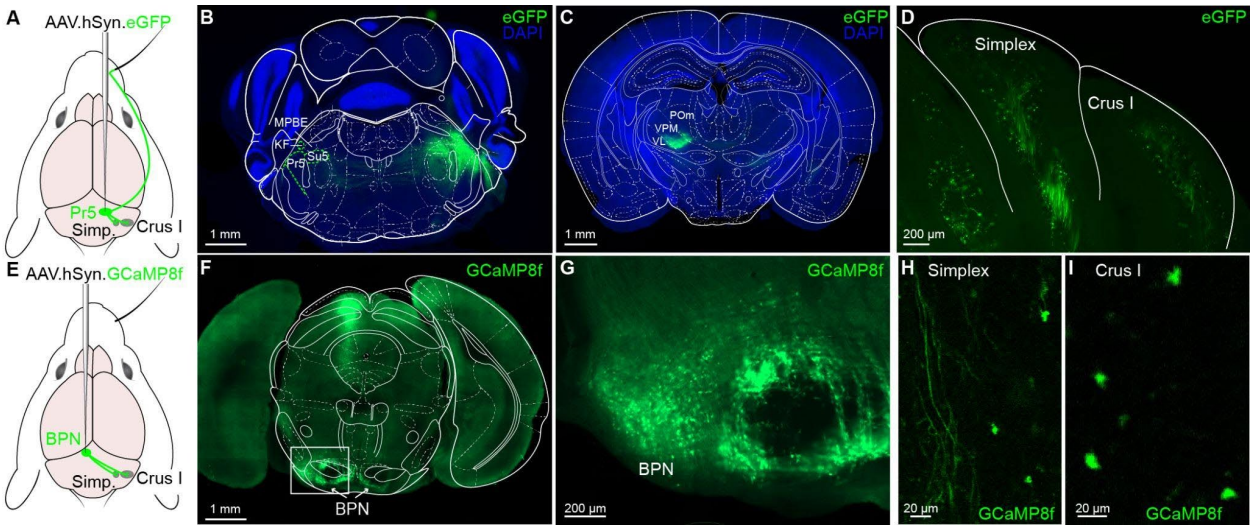

Fig. S3. Viral expression in trigeminal nuclei and pons after AAV injections targeting mossy fiber pathways. (A) Schematic of AAV.hSyn.eGFP injection into the principal trigeminal nucleus (Pr5) and the resulting projection pattern to cerebellar Simplex and Crus I. **(B)** Coronal section showing eGFP-labeled Pr5 projection fibers (green) in the brainstem. DAPI (blue) outlines anatomical landmarks. Labeled structures include KF, Pr5, and Su5. **(C)** Coronal section showing eGFP expression in thalamic relay nuclei following Pr5 injection. Fluorescence is visible mostly in VPM, the main thalamic target of Pr5. DAPI counterstaining is shown in blue. **(D)** eGFP-labeled mossy fiber terminals in the cerebellar cortex following Pr5 injection. Expression is visible in Simplex and Crus I. White outlines mark lobular boundaries. **(E)** Schematic of AAV.hSyn.GCaMP8f injection into the basilar pontine nuclei (BPN) and the resulting cerebellar projections. **(F)** Coronal section showing GCaMP8f expression in the pons after BPN injection. The white box indicates the region shown in (g). **(G)** Higher-magnification view of the boxed region in (f), showing GCaMP8f-expressing neurons and fibers within the basilar pontine nuclei (BPN). **(H,I)** Mossy fiber terminals in the cerebellar cortex, Simplex and Crus I following BPN injection.

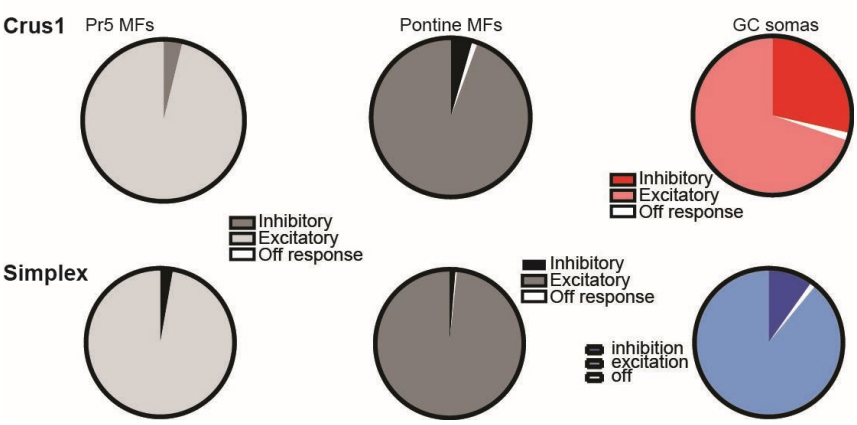

Fig. S4. Response polarity distributions. In Crus I, most responses were excitatory in Pr5 MFs (96%), pontine MFs (94%), and GCs (70%), followed by inhibitory responses (Pr5 MFs, 4%; pontine MFs, 5%; GCs, 29%). Off responses were rare (Pr5 MFs, 0%; pontine MFs, 1%; GCs, 2%). In Simplex, 10% of GC responses were inhibitory, and 1% were off responses. In Simplex pontine MFs, 1% of responses were inhibitory, and 0.5% were off responses, whereas in Pr5 MFs, 3% were inhibitory, and no off responses were observed.

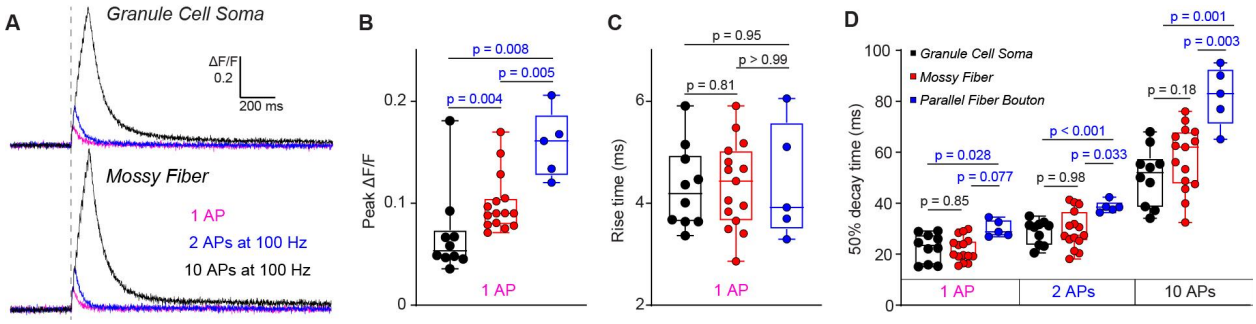

Fig. S5. GCaMP8f calcium signals in granule cell somata, mossy fibers, and parallel fiber boutons. **(A)** Example calcium transients recorded from a granule cell soma (top) and a mossy fiber bouton (bottom) in response to 1 action potential (AP; magenta), 2 APs at 100 Hz (blue), and 10 APs at 100 Hz (black). The ΔF/F scale and time scale are indicated. **(B)** Peak ΔF/F values for granule cell somata (median = 0.05, *n* = 10), mossy fibers (0.09, *n* = 15), and parallel fiber boutons (0.1, *n* = 5) in response to 1 AP. Statistical comparisons (Mann–Whitney tests): soma vs mossy fiber (*P* = 0.004); mossy fiber vs bouton (*P* = 0.005); soma vs bouton (*P* = 0.008). (**C)** Rise times for granule cell somata (median = 4.2 ms, *n* = 10), mossy fibers (4.4 ms, *n* = 15), and parallel fiber boutons (3.9 ms, *n* = 5) in response to 1 AP. Statistical comparisons (Mann–Whitney tests): soma vs mossy fiber (*P* = 0.8); mossy fiber vs bouton (*P* > 0.99); soma vs bouton (*P* = 0.95). **(D)** 50% decay times measured for 1 AP, 2 APs at 100 Hz, and 10 APs at 100 Hz. 1 AP: granule cell somata (median = 23.5 ms, *n* = 10), mossy fibers (19.9 ms, *n* = 15), and parallel fiber boutons (28.8 ms, *n* = 5). Statistical comparisons (Mann–Whitney tests): soma vs mossy fiber (*P* = 0.85); mossy fiber vs bouton (*P* = 0.008); soma vs bouton (*P* = 0.03). 2 APs at 100 Hz: somata (30.9 ms, *n* = 10), mossy fibers (27.4 ms, *n* = 15), and boutons (38.5 ms, *n* = 5). Statistical comparisons: soma vs mossy fiber (*P* = 0.98); mossy fiber vs bouton (*P* = 0.03); soma vs bouton (*P* = 0.0007). 10 APs at 100 Hz: somata (52.1 ms, *n* = 10), mossy fibers (62.1 ms, *n* = 15), boutons (83 ms, *n* = 5). Statistical comparisons: soma vs mossy fiber (*P* = 0.18); mossy fiber vs bouton (*P* = 0.003); soma vs bouton (*P* = 0.001).

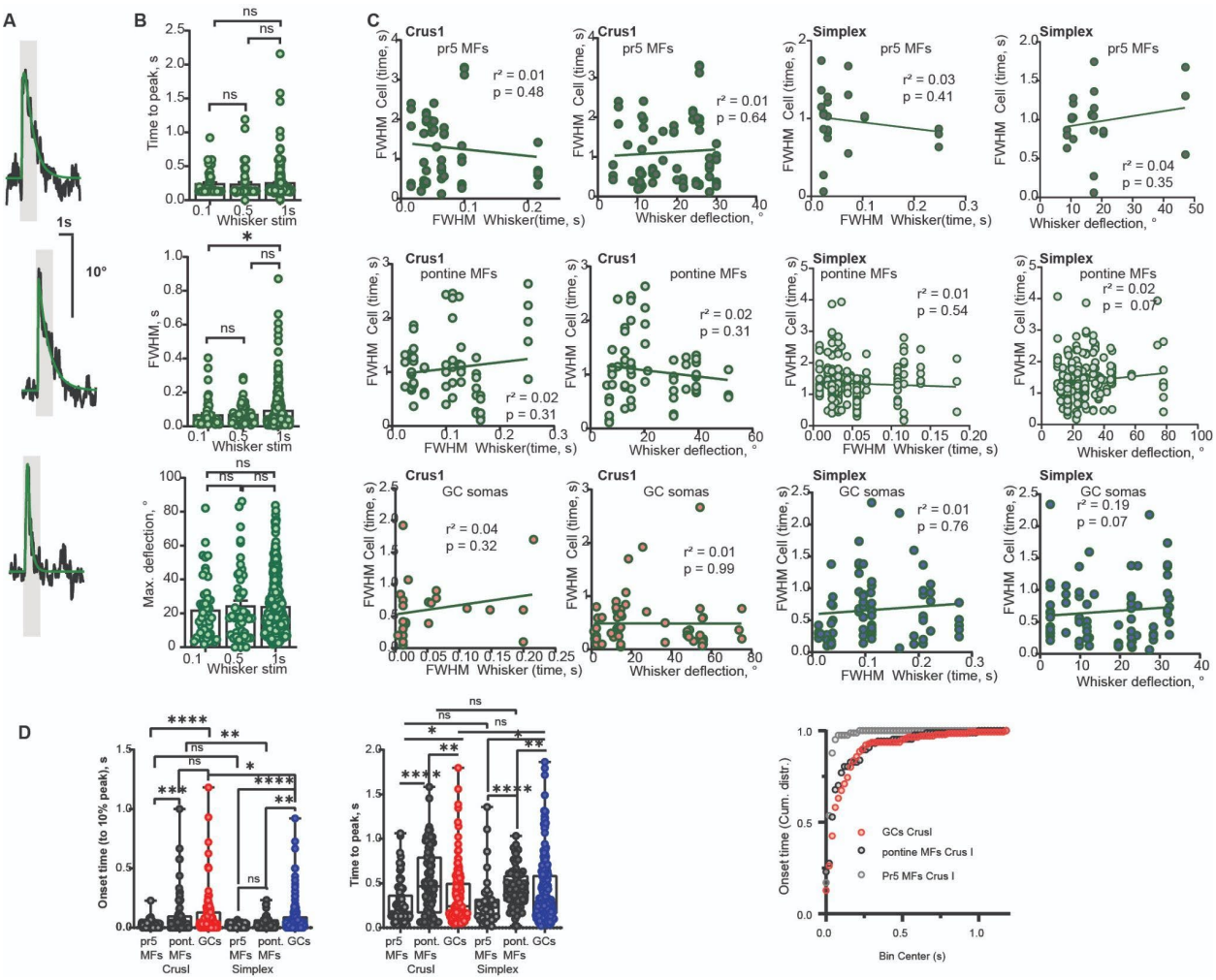

Fig. S6. Whisker kinematics and timing of cerebellar MG and GC responses to whisker air-puff stimulation. (A) Representative curve fits of contralateral whisking response to 1s whisker air puff stimulation, averaged across trials. **(B)** Time to peak of averaged over trials whisker response of the whisker contralateral to the air puff was not different for 1 s (0.25 ± 0.01 ms, *N* = 29 mice, *n* = 379 experiments) compared to 100 ms (0.24 ± 0.02 s, *N* = 17, *n* = 59, *P = 0.6*), or 500 ms (0.23 ± 0.03 s, *N* = 18, *n* = 54, *P = 0.34*) whisker air puff stimulation. FWHM of Whisker response was not different between 0.1 s (0.06 ± 0.01s, *n* = 51, *N* = 17) and 0.5 s (0.07 ± 0.01 s, *n* = 50, *N*= 17, *P = 0.15*) or 0.5 and 1 s (0.09 ± 0.01 s, *n* = 340, *N* = 31, *P = 0.48*), but slightly longer for 1 s compared to 0.1 s (*P = 0.02*). Maximal whisker deflection was not different for 1 s whisker air puff stimulation (24.02 ± 0.93°, *N* = 27, *n* = 330) compared to 100 ms (21.88 ± 2.62°, *N* = 15, *n* = 48, *P = 0.2*) or 500 ms (24.43 ± 3.06°, *N* = 17, *n* = 50, *P = 0.51*). **(C)** FWHM of whisker response did not correlate with FWHM of CrusI Pr5 MF, pontine MF or GC responses. Maximal deflection of whisker response did not correlate with FWHM of CrusI pr5 MF (*N* = 4 mice, *n* = 76 cells), pontine MF (*N* = 3, *n* = 86) or GC (*N* = 5, *n* = 83) responses. FWHM of whisker response did not correlate with FWHM of Simplex GCs, pontine MF, or Pr5. **(D)** Response onset was delayed in Crus I GCs (0.1 ± 0.014 s) compared with Pr5 MFs (0.03 ± 0.005 s, *P* < 0.0001) but not significantly compared with pontine MFs (0.1 ± 0.02 s, *P* = 0.2). Response onset was later in Simplex GCs (0.08 ± 0.009 s) than in pontine MFs (0.04 ± 0.005 s, *P* = 0.0018) or Pr5 MFs (0.02 ± 0.003 s, *P* < 0.0001). Time to peak was later in Crus I pontine MFs (0.51 ± 0.04 s) than in Pr5 MFs (0.27 ± 0.02 s, *P* = 0.0006) or GCs (0.36 ± 0.03 s, *P* = 0.002). Also in Simplex, time to peak was later in pontine MFs (0.43 ± 0.02 s) than in Pr5 MFs (0.29 ± 0.04 s, *P* < 0.0001) or GCs (0.41 ± 0.03 s, *P* = 0.007). Right, cumulative distribution of Onset time for GCs, pontine MFs, and Pr5 MFs. Note the overlap of GCs and pontine MFs.

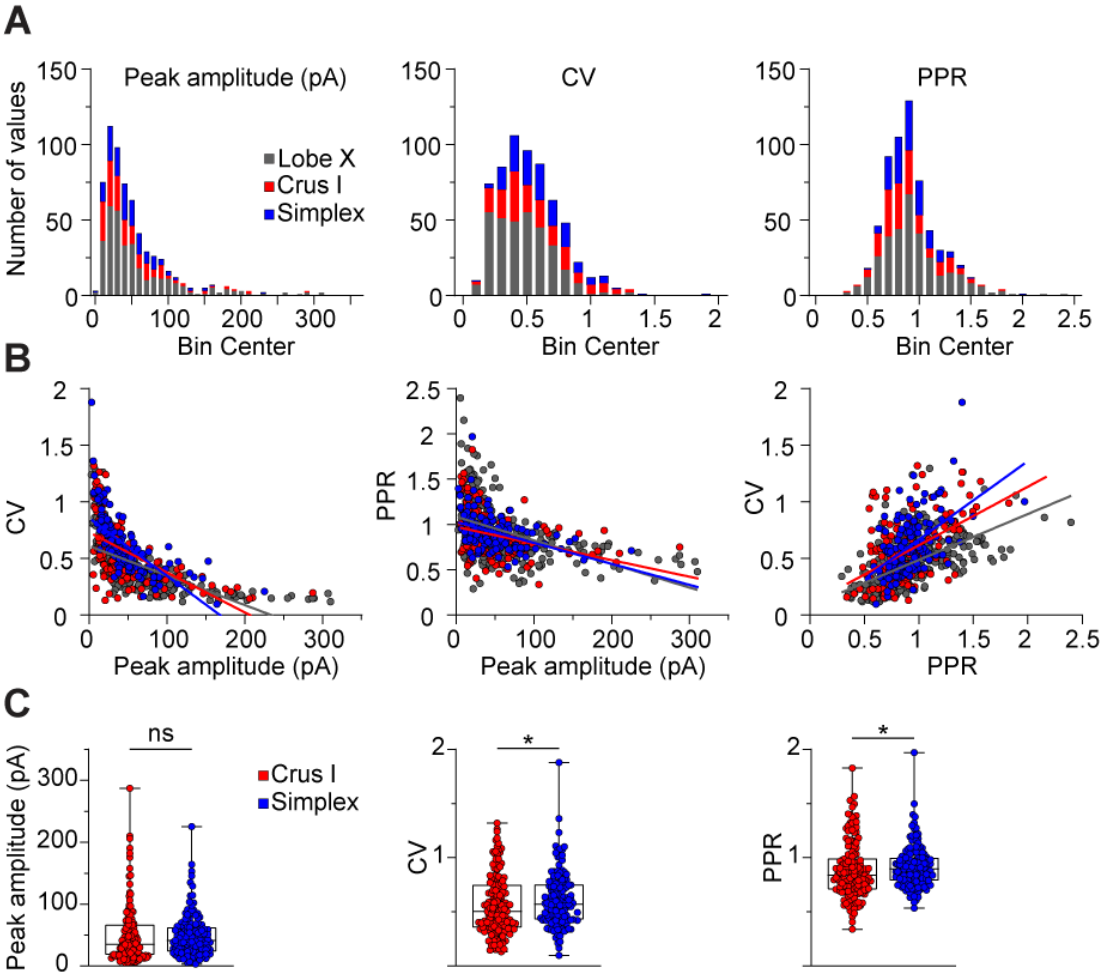

Fig. S7. Diverse MF-GC synaptic response metrics in Crus I and Simplex mirror vestibulocerebellar heterogeneity but vary between regions. (A) Stacked frequency distributions of EPSC peak amplitude (Lobe X: 37.50 ± 55.60, Crus I: 34.97 ± 45.19, Simplex: 41.30 ± 34.17), coefficient of variation of peak amplitude (CV; Lobe X: 0.45 ± 0.21, Crus I: 0.50 ± 0.27, Simplex: 0.57 ± 0.25), and paired-pulse ratio (PPR; Lobe X: 0.89 ± 0.31, Crus I: 0.84 ± 0.26, Simplex: 0.90 ± 0.20) at 100 Hz stimulation from 326 MF-GC synapses recorded in Lobe X (vestibulocerebellum, see Chabrol et al., 2015), 158 synapses in Crus I and 143 synapses in Simplex using a blind MF stimulation protocol. Distribution shapes are similar across areas, suggesting similar diversity in STP dynamics at MF-GC synapses. **(B)** Correlation plots between EPSC peak amplitude, CV and PPR across the recorded synapses in each region, with corresponding regression lines (Peak amplitude vs CV; Lobe X: R^2 = 0.46, Crus I: R^2 = 0.34, Simplex: R^2 = 0.47; Peak amplitude vs PPR; Lobe X: R^2 = 0.22, Crus I: R^2 = 0.10, Simplex: R^2 = 0.16; PPR vs CV; Lobe X: R^2 = 0.34, Crus I: R^2 = 0.23, Simplex: R^2 = 0.29). Significant correlations were detected across all parameters in each region (P < 0.0001 for all pairs, Spearman correlation). (**C**) Quantification of peak amplitude (left; Crus I: 34.97 ± 45.19 pA, Simplex: 41.30 ± 34.17 pA), CV (middle; Crus I: 0.50 ± 0.27, Simplex: 0.57 ± 0.25), and PPR (right; Crus I: 0.84 ± 0.26, Simplex: 0.90 ± 0.20). Simplex synapses show higher CV (P = 0.024, Mann-Whitney test) and PPR (P = 0.024, Mann-Whitney test) (Crus I: n = 158, Simplex: n = 143).

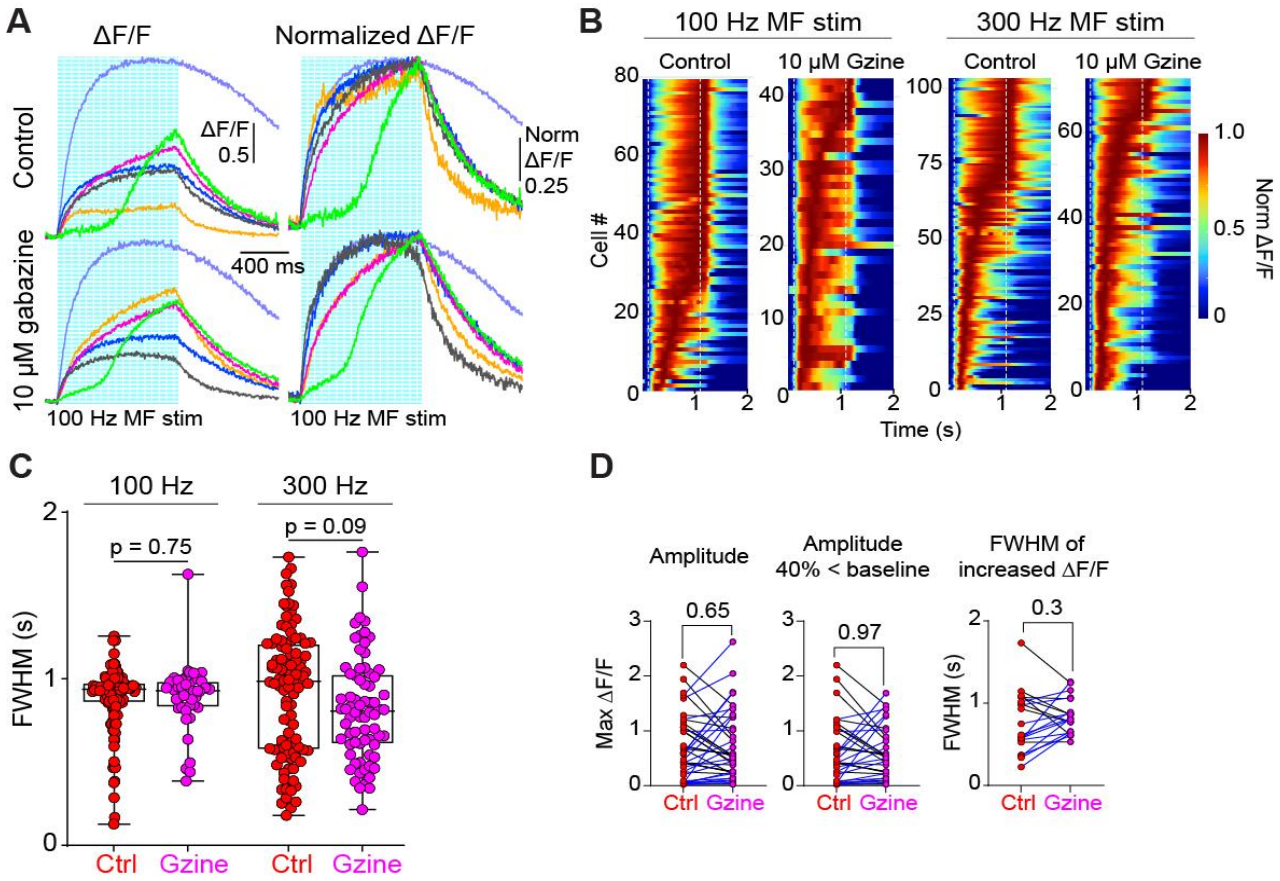

Fig. S8. Effect of GABA$_A$ receptor blockade on GC calcium responses during MF stimulation. (A) Example ΔF/F calcium traces (left) and normalized ΔF/F traces (right) recorded from granule cells during 100 Hz MF stimulation under control conditions (top) and in the presence of 10 μM gabazine (bottom). Individual cells are color-coded. Vertical blue lines indicate the timing of each MF stimulus pulse. **(B)** Raster plots of normalized ΔF/F signals from simultaneously recorded granule cells during 100 Hz (left) and 300 Hz (right) MF stimulation under control conditions and after application of 10 μM gabazine. Each row represents a single cell. Dashed white lines mark the onset and offset of the MF stimulation. Color scale indicates normalized ΔF/F. (**C)** Full width at half maximum (FWHM) of granule cell calcium responses during 100 Hz (left) and 300 Hz (middle) MF stimulation under control conditions and with gabazine. Box plots show the median and interquartile range; each point represents a single cell (100 Hz: control median = 0.94 s, *n* = 117; gabazine median = 0.93 s, *n* = 42. 300 Hz: control median = 0.98 s, *n* = 104; gabazine median = 0.81 s, *n* = 72; Mann–Whitney test *P* = 0.09). **(D)** Summary of granule cell calcium response measurements under control conditions and after application of 10 μM gabazine. The left plot shows the peak ΔF/F amplitude for all paired cells (Wilcoxon matched-pairs signed-rank test, *P* = 0.65; n = 42). The middle plot shows peak ΔF/F amplitude after excluding cells whose baseline fluorescence increased by more than 40% during gabazine application, used as a control to ensure that effects were not driven by baseline shifts (Wilcoxon matched-pairs signed-rank test, *P* = 0.97; n = 35). The right plot shows the FWHM measured only from cells whose ΔF/F increased during gabazine application (Wilcoxon matched-pairs signed-rank test, *P* = 0.33; n = 20).

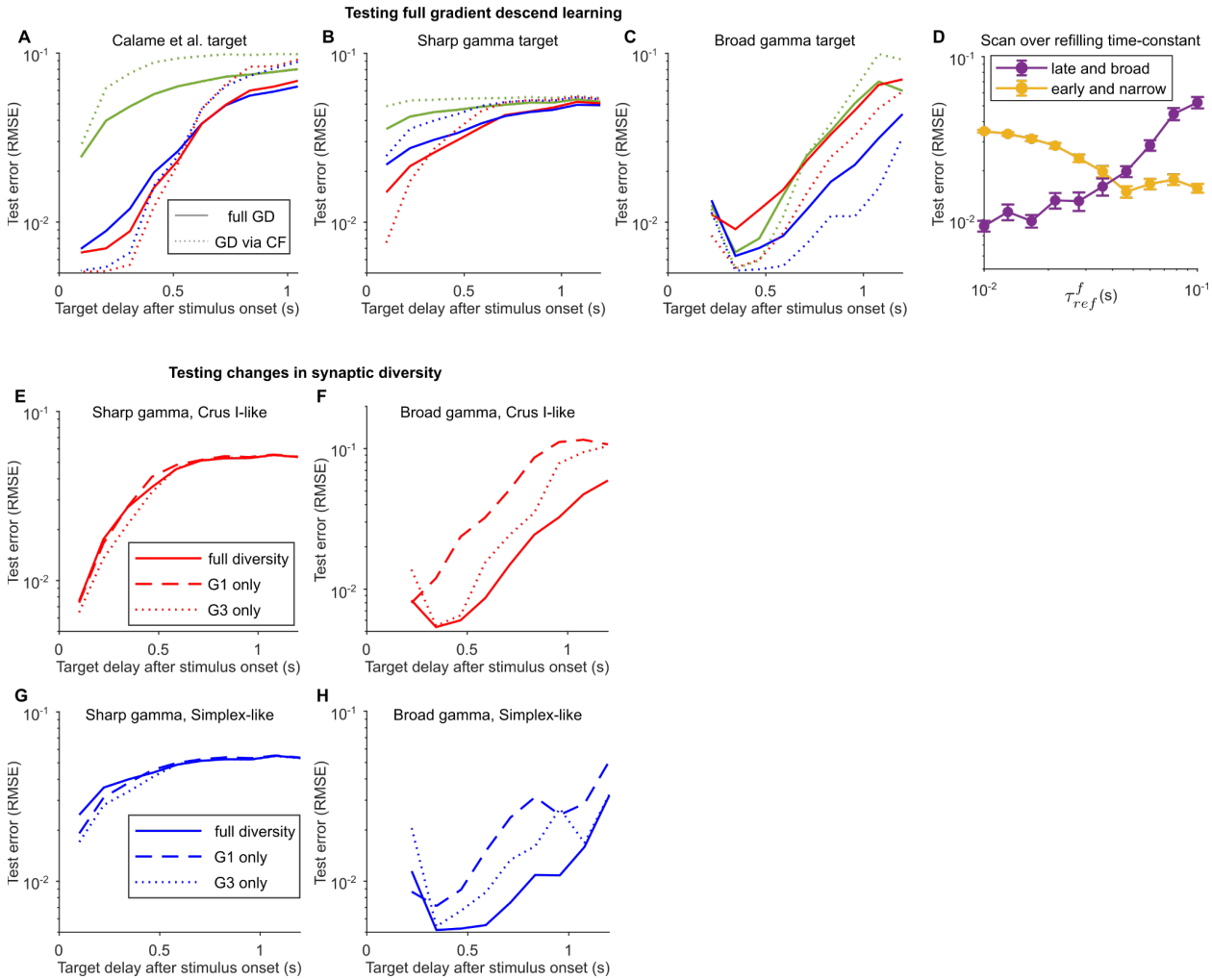

Fig. S9. Full vs restricted gradient descent learning, role of synaptic refilling time-constant, and effect of synaptic diversity on reconstruction error. (A) Reconstruction error of a sharp temporal signal representing forelimb reach kinematics (Calame et al., 2023) at varying target delays for different GC basis sets with biologically inspired learning rule (dotted lines) and full gradient descent (solid lines); each point represents an average over 30 learning simulations. **(B)** Same as (**A**), but for narrow gamma target signals with a SD of 0.01 s. **(C)** Same as (**B**), but for broad gamma target signals with a SD of 0.15 s. **(D)** Reconstruction error of gamma target signals with a delay of 0.25 s and with a SD of 0.01 s (yellow line) and with a delay of 0.75 s and with a SD of 0.15 s (purple line), for varying fast vesicle refilling time-constants. **(E)** Reconstruction error of gamma target signals with a SD of 0.01 s at varying target delays for Crus I-like models; different lines denote models with varying degrees of synaptic diversity: full diversity (solid line), all synapses assigned to high release probability group 1 (dashed line), and all synapses assigned to low release probability group 3 (dotted line). **(F)** Same as (E) but for broad gamma target signals with a SD of 0.15 s. **(G,H)** Same as (**E,F**) but for the Simplex-like model.

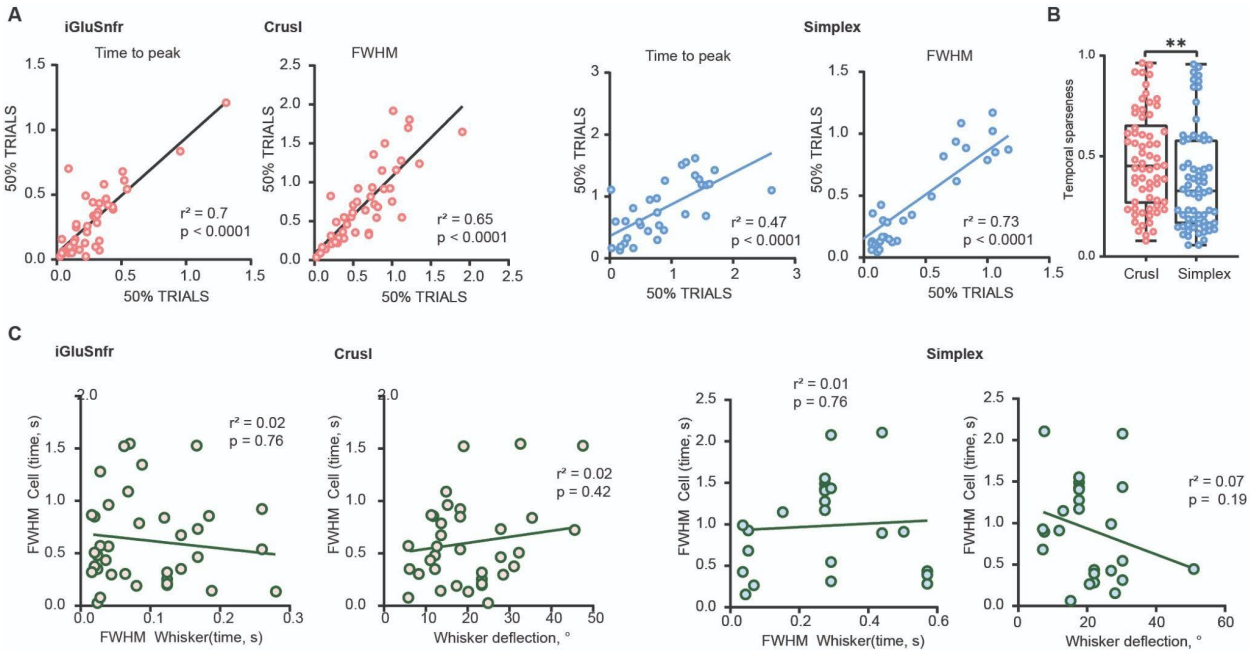

Fig. S10. Reliable and region-specific temporal dynamics of MF–GC synaptic responses to whisker airpuff. (A) GCs were sparsely labelled with iGluSnFR under the TRE promoter in TCGO mice. Correlations calculated from randomly averaged 50% of trials, respectively, for MF–GC synapses in Crus I and Simplex for time to peak and FWHM. **(B)** MF–GC synaptic responses were more temporally sparse in Crus I (0.48 ± 0.03 s) than in Simplex (0.38 ± 0.03 s, P = 0.008). **(C)** No correlation was observed between MF–GC synaptic temporal dynamics and responses to contralateral whisker stimulation in experiments in which the average whisker response could be fitted to a multiexponential function (N = 9, n = 58; Simplex, N = 8, n = 40).