A PLASTICITY MODEL OF CEREBELLAR REGULATION OF CORTICAL DYNAMICAL STABILITY AND TASK-ASSOCIATED REPRESENTATIONS

Anonymous authors
Paper under double-blind review

ABSTRACT

Experiments indicate that cerebellar regulation helps stabilize cortical activity, but the mechanisms that support this effect remain unclear. We construct a biologically motivated cortical–cerebellar model to study how cerebellar learning can regulate an active cortical circuit. A recurrent neural network (RNN) with fixed connections and active/chaotic dynamics represents the cortical circuit, while continuous online learning acquires cerebellar feedback. Differential Hebbian plasticity at the input stage associates cortical pattern activity with category-related cues (context) supplied by an auxiliary cortical population; the output stage uses eligibility traces and a teaching signal q that reflects the history of internal activity changes to learn feedback. In a five-category pattern task, removing context after training leaves cerebellar feedback able to reduce trajectory speed and repeated-stimulus perturbation distance while increasing category separation at the final state. Structural replacement and plasticity ablations show that the expanded sparse feedforward cerebellar pathway and coordinated adaptation at the input and output sites jointly support response stabilization and category organization. When category cues are mismatched, part of the dynamical control is preserved but category separation is impaired; unlearned categories evoke stronger q responses and weaker cumulative perturbation contraction. These results suggest that the cerebellum can learn to regulate effective recurrent drive by sensing and integrating internal cortical changes, while associative plasticity allows stabilization to retain task-related pattern structure.

1 INTRODUCTION

The contribution of the cerebellum to cortical processing is central to understanding cortico–cerebellar computation. Experimental studies show that disrupting cerebellar output can impair sustained cortical preparatory activity (Gao et al., 2018; Chabrol et al., 2019) and weaken temporal coordination between sensory and motor cortex (Popa et al., 2013). In a monkey motor-adaptation task, blocking cerebellar output increased movement variability, increased the dimensionality of cortical preparatory representations, and impaired generalization across targets (Israely et al., 2025); population recordings also reveal coordinated changes in cortical and cerebellar activity during motor learning (Wagner et al., 2019). These findings suggest that the cerebellum supports reliable task representations partly by shaping the evolution of cortical states, beyond correcting behavioral outcomes. They raise two questions: how can cerebellar learning stabilize cortical activity against perturbations while retaining task-relevant pattern differences, and how can such learning be implemented with biologically plausible local signals?

We use an active recurrent neural network (RNN) with autonomous chaotic dynamics to model a cortical circuit that requires feedback stabilization. Brief pattern stimulation initializes the cortical response; after sensory drive is withdrawn, internal activity continues to evolve, and small differences in the initial state can be amplified by recurrence (Sompolinsky et al., 1988). This setup makes the cerebellar contribution explicit as learned control of autonomous cortical dynamics. Existing methods show that chaotic networks can learn reliable temporal outputs through output feedback or training on their own trajectories, including FORCE learning and spontaneous-trajectory training (Sussillo & Abbott, 2009; Laje & Buonomano, 2013); predictive alignment further provides a

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

local online rule for learning recurrent interactions (Asabuki & Clopath, 2025). However, specifying target outputs or activity trajectories supplies an explicit learning direction and requires error supervision. Without such error supervision, learning must derive a control signal from available neural activity while organizing the controlled states into task-related structure. Suppressing perturbations alone does not determine which patterns should be grouped together, and partitions arising from input statistics need not match task categories. Reliable representations therefore require associative experience to organize how different states share control.

We address this problem with a model coupling a fixed cortical RNN to a plastic cerebellar feedforward pathway. Inspired by expansion coding in the granule-cell layer and cerebellar associative learning (Marr, 1969; Albus, 1971; Sendhilnathan et al., 2024), the input stage uses a differential Hebbian rule to associate cortical activity with category-related cues (context) supplied by an auxiliary cortical population, thereby organizing the granule representations recruited by different patterns. The auxiliary cortex denotes a task-related higher-order association population that may provide category- or context-related activity and is functionally distinct from the cortical circuit carrying the main recurrent dynamics. At the output stage, eligibility-trace learning (Gerstner et al., 2018) combines synaptic activity history with an internal-change-sensing signal q to learn feedback that regulates effective cortical recurrence. The teaching role of q is motivated by climbing-fiber responses to novel events and violations of expectation and by their role in associative learning (Ohmae & Medina, 2015; Silva et al., 2024); q itself is a modeling construct for history-dependent sensing of internal activity changes. Category-related information is supplied by associative input, and neither plastic site computes the difference between a target output and the actual output. Control is instead based on activity changes generated within the network, allowing local activity, synaptic traces, and endogenous modulation to support a biologically motivated feedback-learning mechanism.

After training on five pattern classes, we freeze plasticity and withdraw the associative cues. Learned cerebellar feedback still reduces cortical state drift, suppresses amplification of initial-state perturbations for the same pattern, and forms a category-related final-state structure. Structural replacement shows that the tested cerebellar feedforward pathway is more favorable than a cortex-like recurrent feedback module for jointly producing these effects. Interventions on associative input and plasticity further show that partial perturbation suppression can survive damage to category organization, whereas complete representations require coordinated associative experience and feedback learning. The temporal organization of the teaching signal, granule-cell sparsity, and cortical gain jointly shape this coordination; unlearned categories evoke stronger internal-change responses and weaker mean perturbation contraction. Thus, cerebellar learning control improves the consistency of repeated responses while organizing cortical states according to task-related associations.

Our contributions are threefold:

1. We introduce an online cerebellar feedback-learning mechanism driven by internal activity changes. With cortical recurrent connections fixed and the cortical circuit initially in a chaotic regime, the mechanism stabilizes responses without supervised target outputs or errors relative to complete activity trajectories.
2. We show that response stabilization and category organization are coordinated: associative cues jointly shape input representations and output feedback, allowing cerebellar control to retain task structure after the cues are withdrawn.
3. Through structural replacement experiments, we show that the cerebellar feedforward pathway is more favorable under the tested conditions for combining response stabilization with category organization, providing computational evidence for structural division of labor in cortico–cerebellar loops.

2 Related work

Recurrent dynamics and chaos control. Random recurrent networks can generate autonomous chaos (Sompolinsky et al., 1988), whereas reservoir computing uses evolving states to process input history (Maass et al., 2002; Dambre et al., 2012). Rich activity, perturbation amplification, and task relevance are not equivalent: sustained input can suppress chaos, and nonnormal transients can support computation in otherwise stable networks (Rajan et al., 2010; Hennequin et al., 2014). FORCE, trajectory training, and full-FORCE organize unstable activity into reliable temporal outputs (Sus-

sillo & Abbott, 2009; Laje & Buonomano, 2013; DePasquale et al., 2018); local online rules further combine eligibility traces with predictive alignment (Bellec et al., 2020; Asabuki & Clopath, 2025). These approaches typically use target outputs, teacher trajectories, or error signals to specify the learning direction. We place plasticity in the cerebellar pathway, use internal changes as the control basis, and use associative input during training to organize task-related states.

Cortico–cerebellar coupling and cerebellar learning. Pontine afferents and nuclear–thalamic efferents connect the cerebellum with motor and association cortex (Kelly & Strick, 2003; Ramnani, 2006); experiments show that cerebellar output can alter cortical preparatory activity, temporal coordination, and task representations (Gao et al., 2018; Popa et al., 2013; Israely et al., 2025; Wagner et al., 2019). The granule-cell layer integrates multiple inputs and forms separated representations through sparse expansion (Huang et al., 2013; Litwin-Kumar et al., 2017; Cayco-Gajic et al., 2017; Nguyen et al., 2023; Garcia-Garcia et al., 2026); prefrontal category coding and cortico–cerebellar associative circuits provide physiological reference points for the auxiliary-cortex context used here (Freedman et al., 2001; Siegel et al., 2012; Sendhilnathan et al., 2024). Climbing-fiber responses to novel cues and violated expectations, together with graded dendritic actions, provide a biological reference for teaching by internal changes (Ohmae & Medina, 2015; Silva et al., 2024; Zang & De Schutter, 2019). Recent models examine predictive feedback, task acquisition, and complementary cortico–cerebellar computation (Dean et al., 2010; Boven et al., 2023; Pemberton et al., 2024; Bao et al., 2026), but few address stabilization without error supervision together with associative representation. We test this problem with plasticity at two sites and use temporal plasticity and eligibility traces as physiological reference mechanisms (Sgritta et al., 2017; Gao et al., 2012; Suvrathan et al., 2016).

3 Model and Learning Rules

3.1 Cortico–cerebellar model

The model consists of a fixed cortical recurrent network and a plastic cerebellar feedforward pathway (Figure 1). The cortex contains N units with state and activity $\boldsymbol{x}, \boldsymbol{r} \in \mathbb{R}^N$; the cerebellum contains G granule units whose activity $\boldsymbol{h} \in \mathbb{R}^G$ is transformed by downstream pathways into modulation $\boldsymbol{m} \in \mathbb{R}^N$. Indices i, j and k refer to cortical and granule units, respectively. Let J denote fixed recurrent connectivity, $\boldsymbol{F}$ the recurrent drive regulated by the cerebellum, and τ_r the cortical time constant:

$$\tau_r \dot{\boldsymbol{x}} = -\boldsymbol{x} + \boldsymbol{F}(\boldsymbol{r}, \boldsymbol{m}), \qquad \boldsymbol{r} = \tanh \boldsymbol{x}, \qquad \boldsymbol{F}(\boldsymbol{r}, \boldsymbol{0}) = J\boldsymbol{r}. \tag{1}$$

Positive modulation increases effective excitation and decreases inhibition, whereas negative modulation has the opposite effect; the detailed gain construction is given in Appendix A.1. A brief pattern stimulus occurs at $t = 0$. After the sensory drive is withdrawn, activity evolves in the closed loop; learning is carried by the input and output mappings $P, W \in \mathbb{R}^{G \times N}$.

An auxiliary cortical population provides associative input $\boldsymbol{c}_\ell(t)$ for pattern class ℓ. Its M_c units are indexed by a and have preferred angles $\varphi_a = 2\pi a / M_c$; class angles θ_ℓ are a random permutation of equally spaced positions. Let κ_c control tuning width, let f_a denote the activity, and let a fixed nonnegative sparse projection $Q \in \mathbb{R}^{G \times M_c}$ route K_c inputs to each granule unit. With fixed gain γ_c,

$$\begin{aligned} f_a(\theta_\ell) &= \exp\left[\kappa_c\{\cos(\theta_\ell - \varphi_a) - 1\}\right], \\ \boldsymbol{c}_\ell(t) &= \gamma_c Q \boldsymbol{f}(\theta_\ell). \end{aligned} \tag{2}$$

Context is paired with the cortical response during 100–300 ms of training and is zero at other times and during testing. The population tuning produces overlap between classes, while the random permutation separates class identity from ring order; projection sampling and gain calibration are described in Appendix A.2.

Let $\boldsymbol{z} \in \mathbb{R}^G$ be the granule integration state, let τ_z be its time constant, let TopK_K select the K largest activities, and let $\mathbf{1}$ denote the indicator function. Sparse input projection and competition give

$$\tau_z \dot{\boldsymbol{z}} = -\boldsymbol{z} + P\boldsymbol{r} + \boldsymbol{c}_\ell(t), \qquad h_k = \mathbf{1}\{k \in \mathrm{TopK}_K(\boldsymbol{z})\}. \tag{3}$$

The baseline is $\tau_z = 5$ ms. Selection is determined by the pre-competition amplitudes, and the binary code contains K active units and has total activity K. The downstream state $\boldsymbol{u} \in \mathbb{R}^N$

integrates output with $\tau_u = 15$ ms; α and $m_{\max}$ control feedback strength and saturation scale:

$$\tau_u \dot{\boldsymbol{u}} = -\boldsymbol{u} + W^\top \boldsymbol{h}, \qquad \boldsymbol{m} = \alpha m_{\max} \tanh(\boldsymbol{u}/m_{\max}). \tag{4}$$

Both plastic sites use neural activity and its history and do not compute an error between an external target output and the actual output.

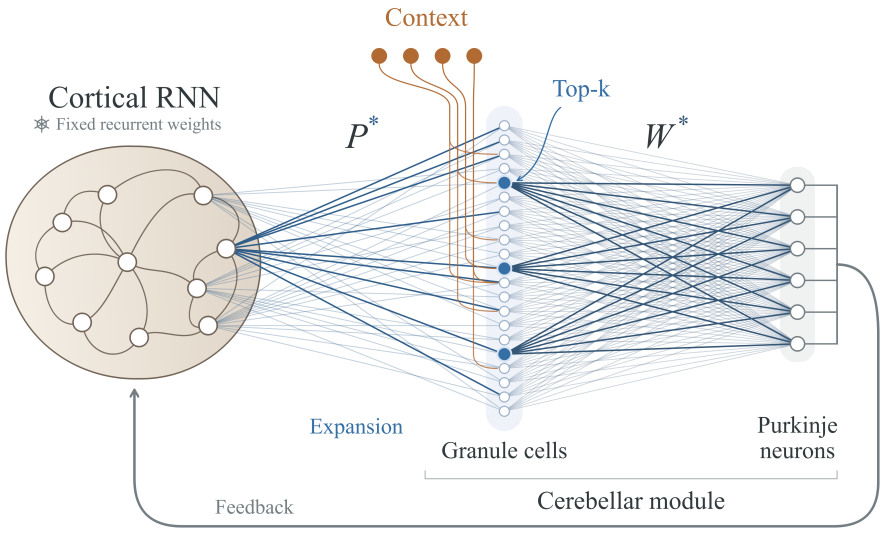


Figure 1: Cortico–cerebellar coupling and plasticity sites. Asterisks mark the trainable P and W mappings, and the snowflake marks fixed cortical recurrent connections. For simplicity, the roles of Purkinje neurons and cerebellar nuclear neurons are merged into an effective output population; the relay from cortex through the pons and mossy fibers to the granule layer is summarized by the input projection, and the return pathway to cortex by the output mapping and filter.

3.2 Temporal associative plasticity at the input site

The input site uses differential Hebbian learning to associate changes in granule activity with recent cortical activity. Let $\mathcal{N}_k$ be the fixed input set of granule unit k, containing F cortical units, and let the initial edge weight be $P^0_{kj} = 1/F$. An allowed set $\mathcal{C}$ specifies the topology: $P_{kj} = 0$ outside the topology, while within the topology $0.1/F \le P_{kj} \le 3/F$ and $\sum_{j\in\mathcal{N}_k} P_{kj} = 1$.

The presynaptic eligibility trace E^P_j starts at zero and retains baseline-shifted activity $1 + r_j$ with $\tau_{E_P} = 100$ ms. Let $\delta = 1$ ms be the neural time step, $\Delta h_k(t) = h_k(t) - h_k(t-\delta)$, and η_P the learning strength. The associative increment on existing connections is

$$\tau_{E_P} \dot{E}^P_j = -E^P_j + (1 + r_j), \qquad \Delta P^{\text{assoc}}_{kj}(t) = \eta_P E^P_j(t)\, \Delta h_k(t). \tag{5}$$

When a granule unit enters or leaves the active set, the associative term strengthens or weakens the input connections that were recently active, pairing changes in population recruitment with cortical response. The forward drive remains $P\boldsymbol{r}$.

P learns during 100–300 ms, with the presynaptic traces accumulating from trial onset. Let $\Delta = 10$ ms be the update interval, let I_Δ be the set of learning sample times within the interval, let P^+ be the updated mapping, and let $\Pi_{\mathcal{C}}$ be Euclidean projection onto $\mathcal{C}$. With decay coefficient λ_P and learning-window length $T_P = 200$ ms, each update is

$$P^+ = \Pi_{\mathcal{C}}\left[P + \sum_{t\in I_\Delta} \Delta P^{\text{assoc}}(t) - \eta_P \lambda_P \frac{\Delta}{T_P}(P - P^0)\right]. \tag{6}$$

The decay term reduces $P - P^0$, and new weights take effect from the next neural step; implementation details are given in Appendix A.3.

3.3 INTERNAL-CHANGE SENSING AND OUTPUT LEARNING

The internal state-deviation vector $\boldsymbol{q} \in \mathbb{R}^N$ senses the history of activity changes and provides teaching information for output learning. With adaptation and gating time constants $\tau_q = 10$ ms and $\tau_\pi = 80$ ms, respectively, $\boldsymbol{q}(0) = \mathbf{0}$ at the beginning of every trial and is updated from activity increments:

$$\boldsymbol{q}(t+\delta) = \left[1 - \frac{\delta}{\tau_q} e^{-t/\tau_\pi}\right] \boldsymbol{q}(t) + \boldsymbol{r}(t+\delta) - \boldsymbol{r}(t). \tag{7}$$

Let the leak gate be $\pi(t) = e^{-t/\tau_\pi}$, giving the continuous form $\dot{\boldsymbol{q}} = \dot{\boldsymbol{r}} - \pi(t)\boldsymbol{q}/\tau_q$. The signal q retains time-weighted net changes, so changes in opposite directions can cancel; early changes decay rapidly, whereas leakage weakens later and prior deviations can continue to influence learning after activity slows. The integral kernel is given in Appendix A.4.

The granule eligibility trace E_k^W retains activity history from trial onset with time constant $\tau_{E_W} = 400$ ms:

$$\tau_{E_W} \dot{E}_k^W = -E_k^W + h_k, \qquad E_k^W(0) = 0. \tag{8}$$

Local susceptibility χ_i weights learning according to the current feedback effect and is determined jointly by the adjustable recurrent drive, cortical response gain, and downstream output saturation (Appendix A.5). W is initialized to zero and updated once per Δ interval during 100–1000 ms. At each update, the $M_W = \min(G, 2K)$ rows with the largest amplitudes in the complete eligibility trace are selected, denoted by $\mathcal{S}(t)$. With learning rate η_W and $\widetilde{W}$ denoting the post-association weights,

$$\widetilde{W}_{ki} = W_{ki} - \eta_W \Delta\, E_k^W q_i \chi_i, \qquad k \in \mathcal{S}(t). \tag{9}$$

With decay rate λ_W, numerical-validity threshold ϵ_E, and weight bound $w_{\max}$, the temporary weights for selected rows with $E_k^W > \epsilon_E$ are multiplied by $1 - \lambda_W \Delta$, after which all selected weights are clipped to $[-w_{\max}, w_{\max}]$; unselected rows retain their previous values. Eligibility traces allow earlier activity to receive later teaching signals; parameters are listed in Appendix A.5.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETTINGS

Pattern stimuli were taken from handwritten digits in the Modified National Institute of Standards and Technology (MNIST) dataset, with digits 0–4 as learned classes and 400 training samples per class, giving 2000 trials. Brief stimulation was implemented by assigning the initial cortical state; sensory input was then zero, and each trial lasted 1 s. The pattern amplitude was scaled to a root-mean-square (RMS) value of 0.5 for each sample, and the standard deviation of the initial-state perturbation was 0.01. The baseline network used $N = 1000$, $G = 32000$, cortical recurrent gain $g = 2$, and a 5% active granule fraction; complete parameters and learning windows are given in Appendix A.

The test set contained independent samples from the same five classes that were not used for training, with 100 samples per class taken from the official MNIST test set. After training, P and W were frozen and context was withdrawn; each sample received two independent initial-state perturbations. The cortex-only control disabled cerebellar feedback while retaining the same recurrent connections and paired stimuli. All conditions used the same sample list and, within each seed, paired input projections and perturbations. Five network seeds (41–45) were used; results were first aggregated with equal weighting across samples and classes and then reported as the mean and sample standard deviation (SD) across seeds. Shading and error bars indicate ± 1 SD. Unlearned classes are examined separately in Section 4.6.

The primary metrics were the five-class mean silhouette coefficient S in the original cortical activity space, trajectory speed $v(t)$, and repeated-stimulus perturbation distance $D(1000)$. S measures category organization, $v(t)$ the rate of activity change, and $D(1000)$ the separation after one second between two independent initial-state perturbations of the same sample. We refer to reduced perturbation amplification and state drift as response stabilization; calculation and averaging order are given in Appendix D, and mechanistic metrics are introduced in their respective result subsections.

4.2 State organization and dynamical control without cues

To test whether learned feedback continues to stabilize cortical responses and preserve category structure after associative input is withdrawn, we compared paired pattern stimulation with the cortex-only condition and with cerebellar feedback retained. Late-period trajectory speed decreased from $(10.163 \pm 1.044) \times 10^{-3}$ ms^{-1} in the cortex-only condition to $(0.434 \pm 0.237) \times 10^{-3}$ ms^{-1} with feedback, while the repeated-stimulus perturbation distance $D(1000)$ decreased from 0.750 ± 0.062 to 0.024 ± 0.014 (Figure 2**a,b**). The five-class mean silhouette coefficient increased from -0.023 ± 0.012 to 0.472 ± 0.024, indicating that reduced perturbation amplification was accompanied by category organization. Principal component analysis (PCA) of final states illustrates this structure (Figure 2**c**); each condition was fitted separately using 30 prespecified samples per class, while quantitative evaluation was performed in the original activity space of the complete test set.

To examine how this stabilization arises, we compared the actual drive with the drive after instantaneous removal of modulation at the same state along the closed-loop trajectory. We divided the RMS difference by the RMS of the unmodulated net drive and by the total effective excitatory and inhibitory drive, yielding R_{net} and R_{total}. At 1000 ms, these quantities were 0.593 ± 0.045 and 0.077 ± 0.005, respectively (Figure 2**d**). Cancellation between excitatory and inhibitory contributions allows a small correction on the scale of total drive to produce a larger change in net drive. The ratio R_{drift} of drift magnitudes at the same state before modulation and after its instantaneous removal was 0.019 ± 0.014 (Figure 2**e**); definitions and the drive decomposition are given in Appendix D.3. This comparison measures the immediate action of feedback and differs from the complete control trajectory obtained by disabling feedback from the stimulation onset. Activity heatmaps further show the reduction in sustained state changes (Figure 2**f**).

4.3 Comparison of cerebellar feedforward and cortical recurrent structures

To further test the role of feedback structure, we compared the complete modeled cerebellar feedforward pathway (FF), including sparse P projection, expansion, Top-K competition, plastic W readout, and output filtering, with a cortex-like recurrent replacement module (RR). At the model level, these two physiologically motivated organizations shared the same cortical substrate and paired training protocol and were configured to have comparable single-step control forward costs; initial input-drive and readout changes were calibrated as described in Appendix E. RR used dense input, continuous hidden activity, and internal recurrence in place of the FF expansion and sparse code; its internal spectral radius ρ_R ranged from 0.5 to 2, while the task-cortex gain remained $g = 2$. FF achieved both higher category separation and lower perturbation distance (Table 1): all RR conditions reduced late-period speed, but the highest silhouette coefficient was 0.072 ± 0.037, below 0.472 ± 0.024 for FF; $D(1000)$ ranged from 0.116 to 0.144, above 0.024 ± 0.014 for FF. This joint advantage is associated with the combination of expansion dimensionality, sparse coding, and feedforward connectivity.

Architecture	Mean silhouette ↑	Late-period speed ↓	$D(1000)$ ↓
Cortex only	-0.023 ± 0.012	10.163 ± 1.044	0.750 ± 0.062
FF	$\mathbf{0.472 \pm 0.024}$	$\mathbf{0.434 \pm 0.237}$	$\mathbf{0.024 \pm 0.014}$
RR, $\rho_R = 0.5$	0.072 ± 0.037	1.583 ± 0.096	0.126 ± 0.019
RR, $\rho_R = 1.0$	0.050 ± 0.031	1.575 ± 0.170	0.116 ± 0.018
RR, $\rho_R = 1.5$	0.048 ± 0.026	1.640 ± 0.130	0.121 ± 0.013
RR, $\rho_R = 2.0$	0.041 ± 0.032	1.812 ± 0.207	0.144 ± 0.025

Table 1: Comparison of feedforward and recurrent feedback modules. Late-period speed is in units of 10^{-3} ms^{-1}; boldface marks the best mean in each column.

4.4 Associative cues and category organization

We examined how category-related input contributes to dynamical control and category organization, considering that context signals may be perturbed by noise in biological systems. We therefore varied cue phase, amplitude, available units, and category correspondence during training (Table 2). For joint perturbations, $\Delta\theta \sim U[-\delta_\theta, \delta_\theta]$ and $A \sim U[1-a, 1+a]$ were sampled on each trial and held fixed within the cue window. Small fluctuations produced little change in the coupled model;

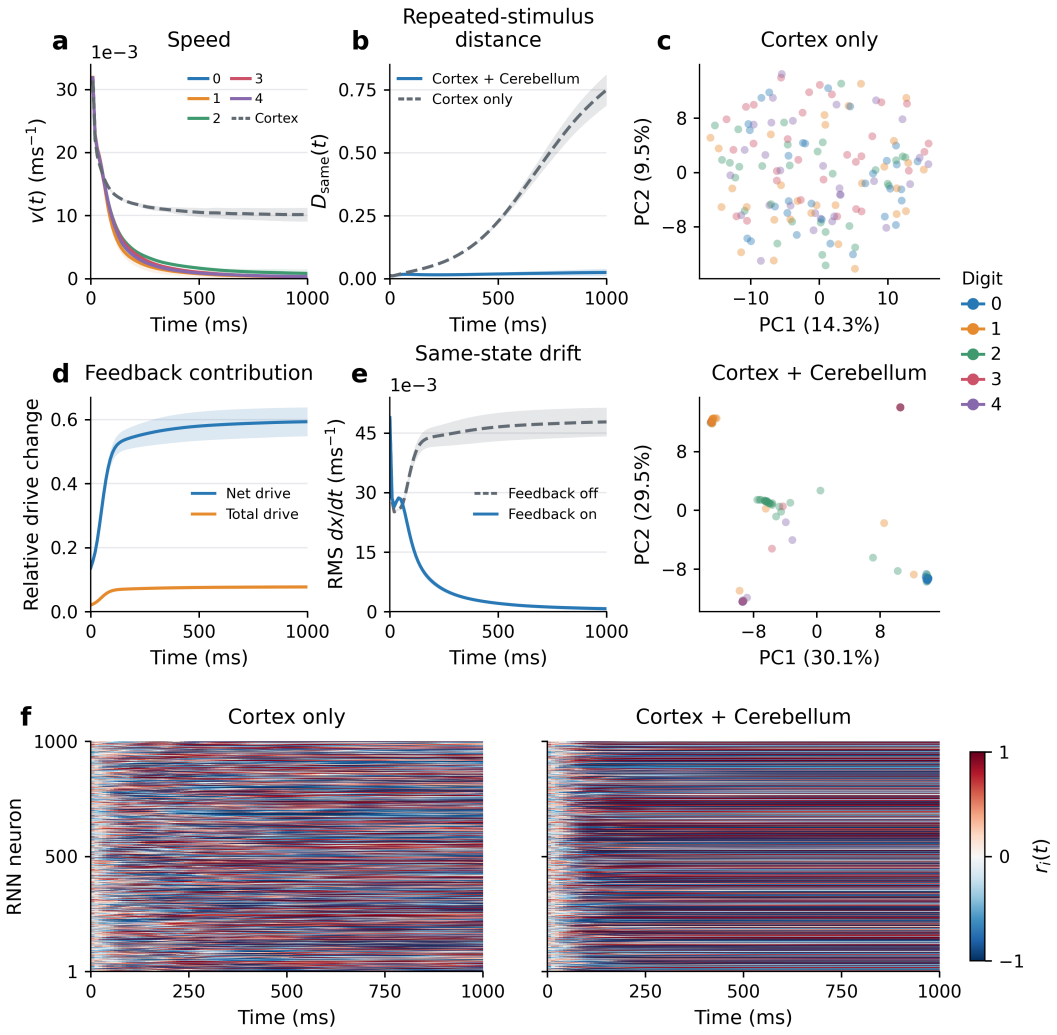


Figure 2: Cerebellar feedback stabilizes cortical responses and organizes category representations. **a**, trajectory speed for each class; the gray dashed line is the five-class mean in the cortex-only condition. **b**, repeated-stimulus perturbation distance. **c**, final-state PCA for seed 45. **d**, feedback-induced changes relative to net and total drive. **e**, drift magnitude with modulation and after instantaneous removal of modulation at the same closed-loop state. **f**, cortical activity for the same prespecified sample in seed 41, with cortex-only on the left and feedback retained on the right. **b,d,e** show equally weighted means across the five classes.

increasing the variation to $(\delta_\theta, a) = (15°, 0.3)$ reduced the silhouette coefficient to 0.268 ± 0.159 while the perturbation distance remained close to baseline. Removing 16 of the 32 auxiliary cortical units on each trial without compensating the remaining input reduced both category separation and repetition consistency. Resampling a cue uniformly from the five classes further changed the pattern–cue correspondence; because resampling could select the original class, the expected mismatch proportion was $0.8p$. At $p = 1$, cues were independent of pattern class, the silhouette coefficient decreased to -0.071 ± 0.019, and $D(1000) = 0.069 \pm 0.020$ remained below the cortex-only condition. Thus, category organization remained relatively robust to small cue noise but was disrupted by substantial mismatch, whereas dynamical control could be maintained through internal teaching at W.

4.5 Multisite plasticity and the teaching signal

To test whether input representations and output feedback must adapt jointly, we separately froze P or W and cross-combined the resulting mappings. With the initial projection P_0 fixed and only W trained, the silhouette coefficient decreased to 0.054 ± 0.050 and $D(1000)$ increased to 0.085 ± 0.023:

Training-cue condition	Mean silhouette	$D(1000)$
Unperturbed	0.472 ± 0.024	0.024 ± 0.014
Joint perturbation, $(\delta_\theta, a) = (5^\circ, 0.1)$	0.437 ± 0.062	0.023 ± 0.012
Joint perturbation, $(\delta_\theta, a) = (15^\circ, 0.3)$	0.268 ± 0.159	0.024 ± 0.005
50% auxiliary-cortex unit dropout	0.114 ± 0.052	0.080 ± 0.032
Class-cue resampling, $p = 0.5$	0.047 ± 0.062	0.061 ± 0.012
Class-cue resampling, $p = 1$	-0.071 ± 0.019	0.069 ± 0.020

Table 2: Effects of variability in associative input during training on uncued performance. Joint perturbations sample phase and amplitude per trial and hold them fixed within the window; unit dropout does not compensate for the remaining input; cue resampling can select the original class, giving an expected mismatch proportion of $0.8p$.

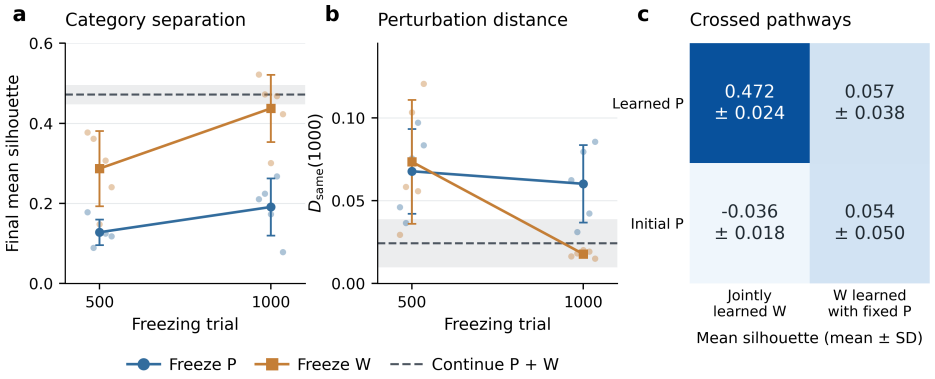


Figure 3: Stage-specific freezing and cross-combination at the two plasticity sites. **a**, final silhouette coefficients after freezing P or W. **b**, corresponding perturbation distances. Freezing occurred at trial 500 or 1000, while the other site continued to trial 2000; gray lines and shading show the mean and SD for joint learning. **c**, cross-combinations of learned P or initial P_0 with the two W mappings; no additional adaptation followed cross-combination.

output learning retained part of perturbation suppression but did not form the category structure obtained by joint learning. To distinguish the contributions of the two sites during training, we froze P or W at a shared checkpoint after trial 500 or 1000 and continued learning at the other site through trial 2000 with paired pattern and perturbation sequences. The final silhouette coefficients after freezing P were 0.128 ± 0.032 and 0.191 ± 0.071, respectively; after freezing W they were 0.287 ± 0.094 and 0.437 ± 0.084 (Figure 3**a**). Freezing W later also retained a low perturbation distance of 0.018 ± 0.002 (Figure 3**b**), indicating that established output feedback can support subsequent input-side adaptation.

Cross-combining the input and output mappings further tested their compatibility. Combining the jointly learned (P_A, W_A) with (P_0, W_B) from W-only training reduced the silhouette coefficients of (P_A, W_B) and (P_0, W_A) to 0.057 ± 0.038 and -0.036 ± 0.018, respectively (Figure 3**c**). Cross-combination loss and stage-specific freezing jointly support coordinated, stage-dependent adaptation at the two sites; training dynamics and the interaction measure for continued learning are provided in Appendix B.

To distinguish the information requirements for feedback control and category organization, we replaced q with the instantaneous activity derivative after calibrating the output-learning strength. The derivative condition restored perturbation suppression ($D(1000) = 0.030 \pm 0.011$), but its silhouette coefficient was only 0.187 ± 0.100 (Appendix E.3). The instantaneous rate decreases as activity slows, whereas q retains the influence of preceding changes; together with the granule eligibility trace, this history dependence may allow activity organization formed during associative input to continue contributing to feedback adaptation.

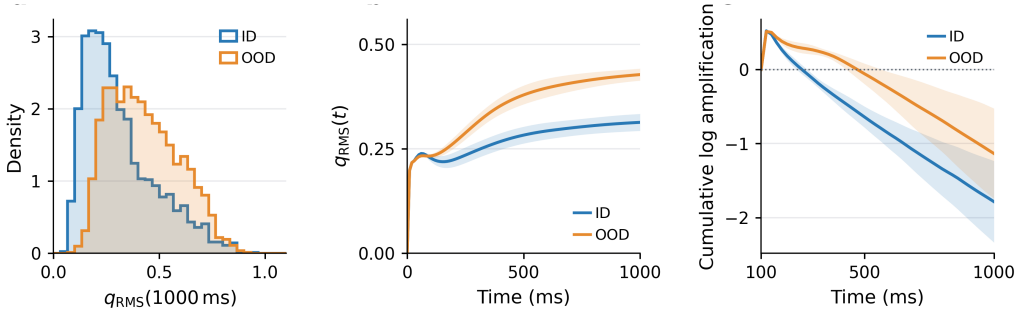


Figure 4: Responses of learned and unlearned classes. **a**, probability density of final-time single-sample q_{RMS} scores, averaged after within-seed normalization. **b**, temporal evolution of the internal-change response. **c**, cumulative logarithmic perturbation gain $L_\epsilon(t)$ for the full closed loop; negative values indicate cumulative contraction along the sampled directions, with $\epsilon = 0.003$.

4.6 Responses to unlearned classes

Beyond the preceding analyses, we examined how the model responds to patterns from unlearned classes. Using the same frozen networks, we compared learned (in-distribution, ID) classes 0–4 with unlearned (out-of-distribution, OOD) classes 5–9. Unlearned classes evoked stronger internal-change responses and weaker mean cumulative contraction. We quantified internal-change sensing with $q_{\mathrm{RMS}}(t) = \mathrm{RMS}[\boldsymbol{q}(t)]$, using the mean across two repetitions at the final time as the image-level score. The ID and OOD final-time responses were 0.313 ± 0.020 and 0.428 ± 0.014, respectively, and the area under the receiver operating characteristic curve (AUROC) was 0.706 ± 0.035 (Figure 4**a**). Differences between groups remained after activity had slowed, reflecting the history dependence of q (Figure 4**b**).

At 100 ms, we applied a perturbation of amplitude $\epsilon = 0.003$ to the full closed-loop state (x, z, u) and periodically renormalized it. The cumulative local logarithmic stretching gave $L_\epsilon(t)$ (Appendix D.5). Values of $L_\epsilon(1000)$ were -1.784 ± 0.550 for ID and -1.135 ± 0.608 for OOD, and the ID mean curve entered the contraction regime earlier (Figure 4**c**). This quantity describes cumulative contraction along sampled finite-amplitude directions and complements the activity-change history recorded by q.

5 Discussion

Our model shows that cerebellar feedback can stabilize responses and support task-related representations in a fixed cortical recurrent circuit. Internal activity changes provide the control basis, while associative input organizes the states that are regulated. Unlike recurrent training based on specified target trajectories (Sussillo & Abbott, 2009; Laje & Buonomano, 2013), feedback changes effective excitatory and inhibitory drive, reducing drift and perturbation amplification while preserving category differences. Under the tested conditions, the calibrated feedforward pathway was superior in both category separation and repetition consistency, suggesting selective rather than uniform activity regulation.

Stabilization and category organization require complementary information. Associative mismatch preserved part of perturbation suppression, whereas category consistency depended on the auxiliary cortex. P/W freezing, cross-combination, and derivative controls show that the two sites and teaching history jointly determine the result: as leakage weakens, q retains preceding changes and eligibility traces retain pathway activity, allowing transient associations to affect later learning. Unlearned classes produced stronger q responses and weaker contraction, consistent with enhanced responses to novel stimuli (Homann et al., 2022), altered population timescales (Garrett et al., 2020), and climbing-fiber responses (Ohmae & Medina, 2015). Top-K and gain scans indicate an operating range in which stabilization and category separation need not improve together (Appendix C).

Effective projections and filters, fixed cortical connectivity, preformed associations, and brief stimulation make this a circuit-level mechanism rather than a complete physiological implementation. Future work should test explicit cell populations, nuclei, and local feedback and teaching signals.

AI USE STATEMENT

Generative AI tools were used to assist with language editing and software development, including suggestions for code implementation and debugging. All AI-generated or AI-modified code was manually reviewed and tested by the authors. Generative AI tools were not used to formulate the main research hypotheses or interpret the experimental results. The authors take full responsibility for the final content of this work.

ETHICS STATEMENT

This study uses publicly available benchmark data and computational models only; it does not involve human or animal subjects.

REPRODUCIBILITY STATEMENT

The model equations, learning rules, complete parameter settings, random seeds, data splits, and evaluation procedures are specified in the main text and Appendix. These details define the computational protocol used to obtain the reported results.

REFERENCES

James S. Albus. A theory of cerebellar function. *Mathematical Biosciences*, 10(1–2):25–61, 1971. doi: 10.1016/0025-5564(71)90051-4.

Toshitake Asabuki and Claudia Clopath. Taming the chaos gently: a predictive alignment learning rule in recurrent neural networks. *Nature Communications*, 16:6784, 2025. doi: 10.1038/s41467-025-61309-9.

Baktash Babadi and Haim Sompolinsky. Sparseness and expansion in sensory representations. *Neuron*, 83(5):1213–1226, 2014. doi: 10.1016/j.neuron.2014.07.035.

Yeyao Bao, Liao Yu, Liangfu Lu, Zhuoqin Yang, and Yunliang Zang. Cerebellar microcircuits enable robust evidence-based decisions through cortico–cerebellar coupling. *Proceedings of the National Academy of Sciences*, 123(35):e2616911123, 2026. doi: 10.1073/pnas.2616911123.

Guillaume Bellec, Franz Scherr, Anand Subramoney, Elias Hajek, Darjan Salaj, Robert Legenstein, and Wolfgang Maass. A solution to the learning dilemma for recurrent networks of spiking neurons. *Nature Communications*, 11:3625, 2020. doi: 10.1038/s41467-020-17236-y.

Ellen Boven, Joseph Pemberton, Paul Chadderton, Richard Apps, and Rui Ponte Costa. Cerebro-cerebellar networks facilitate learning through feedback decoupling. *Nature Communications*, 14:51, 2023. doi: 10.1038/s41467-022-35658-8.

N. Alex Cayco-Gajic and R. Angus Silver. Re-evaluating circuit mechanisms underlying pattern separation. *Neuron*, 101(4):584–602, 2019. doi: 10.1016/j.neuron.2019.01.044.

N. Alex Cayco-Gajic, Claudia Clopath, and R. Angus Silver. Sparse synaptic connectivity is required for decorrelation and pattern separation in feedforward networks. *Nature Communications*, 8: 1116, 2017. doi: 10.1038/s41467-017-01109-y.

François P. Chabrol, Antonin Blot, and Thomas D. Mrsic-Flogel. Cerebellar contribution to preparatory activity in motor neocortex. *Neuron*, 103(3):506–519.e4, 2019. doi: 10.1016/j.neuron.2019.05.022.

Joni Dambre, David Verstraeten, Benjamin Schrauwen, and Serge Massar. Information processing capacity of dynamical systems. *Scientific Reports*, 2:514, 2012. doi: 10.1038/srep00514.

Paul Dean, John Porrill, Carl-Fredrik Ekerot, and Henrik Jörntell. The cerebellar microcircuit as an adaptive filter: Experimental and computational evidence. *Nature Reviews Neuroscience*, 11(1): 30–43, 2010. doi: 10.1038/nrn2756.

Brian DePasquale, Christopher J. Cueva, Kanaka Rajan, G. Sean Escola, and L. F. Abbott. full-FORCE: A target-based method for training recurrent networks. *PLOS ONE*, 13(2):e0191527, 2018. doi: 10.1371/journal.pone.0191527.

David J. Freedman, Maximilian Riesenhuber, Tomaso Poggio, and Earl K. Miller. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, 291(5502):312–316, 2001. doi: 10.1126/science.291.5502.312.

Zhenyu Gao, Boeke J. van Beugen, and Chris I. De Zeeuw. Distributed synergistic plasticity and cerebellar learning. *Nature Reviews Neuroscience*, 13(9):619–635, 2012. doi: 10.1038/nrn3312.

Zhenyu Gao, Courtney Davis, Alyse M. Thomas, Michael N. Economo, Amada M. Abrego, Karel Svoboda, Chris I. De Zeeuw, and Nuo Li. A cortico-cerebellar loop for motor planning. *Nature*, 563:113–116, 2018. doi: 10.1038/s41586-018-0633-x.

Martha G. Garcia-Garcia, Michał J. Wójcik, Srijan Thota, Luke Drake, Amma Otchere, Oluwatobi Akinwale, Lizmaylin Ramos, Rui Ponte Costa, and Mark J. Wagner. Granule cells reorient cortical trajectories to separate contexts. *Nature*, 2026. doi: 10.1038/s41586-026-10946-1.

Marina Garrett, Sahar Manavi, Kate Roll, Douglas R. Ollerenshaw, Peter A. Groblewski, Nicholas D. Ponvert, Justin T. Kiggins, Linzy Casal, Kyla Mace, Ali Williford, Arielle Leon, Xiaoxuan Jia, Peter Ledochowitsch, Michael A. Buice, Wayne Wakeman, Stefan Mihalas, and Shawn R. Olsen. Experience shapes activity dynamics and stimulus coding of VIP inhibitory cells. *eLife*, 9:e50340, 2020. doi: 10.7554/eLife.50340.

Wulfram Gerstner, Marco Lehmann, Vasiliki Liakoni, Dane Corneil, and Johanni Brea. Eligibility traces and plasticity on behavioral time scales: Experimental support of NeoHebbian three-factor learning rules. *Frontiers in Neural Circuits*, 12:53, 2018. doi: 10.3389/fncir.2018.00053.

Guillaume Hennequin, Tim P. Vogels, and Wulfram Gerstner. Optimal control of transient dynamics in balanced networks supports generation of complex movements. *Neuron*, 82(6):1394–1406, 2014. doi: 10.1016/j.neuron.2014.04.045.

Jan Homann, Sue Ann Koay, Kevin S. Chen, David W. Tank, and Michael J. Berry, II. Novel stimuli evoke excess activity in the mouse primary visual cortex. *Proceedings of the National Academy of Sciences*, 119(5):e2108882119, 2022. doi: 10.1073/pnas.2108882119.

Cheng-Chiu Huang, Ken Sugino, Yasuyuki Shima, Caiying Guo, Suxia Bai, Brett D. Mensh, Sacha B. Nelson, and Adam W. Hantman. Convergence of pontine and proprioceptive streams onto multimodal cerebellar granule cells. *eLife*, 2:e00400, 2013. doi: 10.7554/eLife.00400.

Sharon Israely, Hugo Ninou, Ori Rajchert, Lee Elmaleh, Ran Harel, Firas Mawase, Jonathan Kadmon, and Yifat Prut. Cerebellar output shapes cortical preparatory activity during motor adaptation. *Nature Communications*, 16(1):2574, 2025. doi: 10.1038/s41467-025-57832-4.

Roberta M. Kelly and Peter L. Strick. Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of Neuroscience*, 23(23):8432–8444, 2003. doi: 10.1523/JNEUROSCI.23-23-08432.2003.

Rodrigo Laje and Dean V. Buonomano. Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nature Neuroscience*, 16(7):925–933, 2013. doi: 10.1038/nn.3405.

Ashok Litwin-Kumar, Kameron Decker Harris, Richard Axel, Haim Sompolinsky, and L. F. Abbott. Optimal degrees of synaptic connectivity. *Neuron*, 93(5):1153–1164.e7, 2017. doi: 10.1016/j.neuron.2017.01.030.

Wolfgang Maass, Thomas Natschläger, and Henry Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14(11):2531–2560, 2002. doi: 10.1162/089976602760407955.

David Marr. A theory of cerebellar cortex. *Journal of Physiology*, 202(2):437–470, 1969. doi: 10.1113/jphysiol.1969.sp008820.

Tri M. Nguyen, Logan A. Thomas, Jeff L. Rhoades, Ilaria Ricchi, Xintong Cindy Yuan, Arlo Sheridan, David G. C. Hildebrand, Jan Funke, Wade G. Regehr, and Wei-Chung Allen Lee. Structured cerebellar connectivity supports resilient pattern separation. *Nature*, 613:543–549, 2023. doi: 10.1038/s41586-022-05471-w.

Shogo Ohmae and Javier F. Medina. Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. *Nature Neuroscience*, 18:1798–1803, 2015. doi: 10.1038/nn.4167.

Joseph Pemberton, Paul Chadderton, and Rui Ponte Costa. Cerebellar-driven cortical dynamics can enable task acquisition, switching and consolidation. *Nature Communications*, 15:10913, 2024. doi: 10.1038/s41467-024-55315-6.

Daniela Popa, Maria Spolidoro, Rémi D. Proville, Nicolas Guyon, Lucile Belliveau, and Clément Léna. Functional role of the cerebellum in gamma-band synchronization of the sensory and motor cortices. *Journal of Neuroscience*, 33(15):6552–6556, 2013. doi: 10.1523/JNEUROSCI.5521-12.2013.

Kanaka Rajan, L. F. Abbott, and Haim Sompolinsky. Stimulus-dependent suppression of chaos in recurrent neural networks. *Physical Review E*, 82(1):011903, 2010. doi: 10.1103/PhysRevE.82.011903.

Narender Ramnani. The primate cortico-cerebellar system: Anatomy and function. *Nature Reviews Neuroscience*, 7(7):511–522, 2006. doi: 10.1038/nrn1953.

Naveen Sendhilnathan, Andreea C. Bostan, Peter L. Strick, and Michael E. Goldberg. A cerebro-cerebellar network for learning visuomotor associations. *Nature Communications*, 15:2519, 2024. doi: 10.1038/s41467-024-46281-0.

Martina Sgritta, Francesca Locatelli, Teresa Soda, Francesca Prestori, and Egidio Ugo D'Angelo. Hebbian spike-timing dependent plasticity at the cerebellar input stage. *Journal of Neuroscience*, 37(11):2809–2823, 2017. doi: 10.1523/JNEUROSCI.2079-16.2016.

Jennifer J. Siegel, Brian Kalmbach, Raymond A. Chitwood, and Michael D. Mauk. Persistent activity in a cortical-to-subcortical circuit: bridging the temporal gap in trace eyelid conditioning. *Journal of Neurophysiology*, 107(1):50–64, 2012. doi: 10.1152/jn.00689.2011.

N. Tatiana Silva, Jorge Ramírez-Buriticá, Dominique L. Pritchett, and Megan R. Carey. Climbing fibers provide essential instructive signals for associative learning. *Nature Neuroscience*, 27(5):940–951, 2024. doi: 10.1038/s41593-024-01594-7.

H. Sompolinsky, A. Crisanti, and H. J. Sommers. Chaos in random neural networks. *Physical Review Letters*, 61(3):259–262, 1988. doi: 10.1103/PhysRevLett.61.259.

David Sussillo and L. F. Abbott. Generating coherent patterns of activity from chaotic neural networks. *Neuron*, 63(4):544–557, 2009. doi: 10.1016/j.neuron.2009.07.018.

Aparna Suvrathan, Hannah L. Payne, and Jennifer L. Raymond. Timing rules for synaptic plasticity matched to behavioral function. *Neuron*, 92(5):959–967, 2016. doi: 10.1016/j.neuron.2016.10.022.

Mark J. Wagner, Tony Hyun Kim, Jonathan Kadmon, Nghia D. Nguyen, Surya Ganguli, Mark J. Schnitzer, and Liqun Luo. Shared cortex-cerebellum dynamics in the execution and learning of a motor task. *Cell*, 177(3):669–682.e24, 2019. doi: 10.1016/j.cell.2019.02.019.

Yunliang Zang and Erik De Schutter. Climbing fibers provide graded error signals in cerebellar learning. *Frontiers in Systems Neuroscience*, 13:46, 2019. doi: 10.3389/fnsys.2019.00046.

A NUMERICAL IMPLEMENTATION AND PARAMETERS

Table 3 gives the complete settings of the baseline model; differences in the sweep and intervention conditions are described in the main text. Cortical states are integrated with forward Euler, whereas granule-cell states, outputs, and eligibility traces use discrete exponential coefficients for first-order low-pass filtering. Candidate increments of P accumulate on a fixed sparse topology, are submitted every 10 ms, and are projected onto the connectivity constraints defined in Section 3.2; the updated weights participate in the dynamics from the next neural step. For W, associative updates, decay, and amplitude clipping are applied sequentially to the selected eligibility rows. Presynaptic traces are reset to zero before each trial, and the first postsynaptic observation serves as the differencing origin.

Component	Baseline settings
Cortical recurrent network	$N = 1000$; connection density 0.1; gain $g = 2$; $\tau_r = 20$ ms; $r = \tanh(x)$.
Pattern stimuli and initial states	400 training patterns per class, 2000 in total; 100 test patterns per class, with two repetitions; fixed background and Gaussian input projection; per-pattern RMS of 0.5; initial-state noise SD of 0.01; starting-point perturbations sampled independently across trials, with no process noise.
Time discretization	Trial duration 1000 ms; neural time step 1 ms; P/W submission interval 10 ms; evaluation curves recorded every 10 ms.
Granule-cell layer	$G = 32000$, with $F = 16$ fixed sparse inputs per unit; $\tau_z = 5$ ms; Top-$K = 5\%$ ($K = 1600$); winner output $h_k = 1$ and all other outputs zero, without sum normalization.
P initialization and constraints	Fixed nonnegative topology, initial edge weight $1/F$; each row sums to 1; individual edge weights bounded between $0.1/F$ and $3/F$.
P differential learning	$\eta_P = 0.003$; presynaptic signal $1 + r$; presynaptic trace 100 ms; postsynaptic signal given by the actual h, without additional low-pass filtering; window 100–300 ms.
P decay term	$\lambda_P = 0.01$; decay coefficient for the weight displacement at each update $\eta_P \lambda_P \Delta / (200\ \text{ms})$.
W initialization and learning	Initialized to zero; $\eta_W = 1.3 \times 10^{-7}\ \text{ms}^{-1}$; window 100–1000 ms; eligibility-trace time constant 400 ms, without eligibility-amplitude normalization.
W updates and constraints	Complete eligibility traces retained, with the $2K = 3200$ rows of highest eligibility updated at each submission; decay rate $10^{-7}\ \text{ms}^{-1}$; weights clipped to $[-0.01, 0.01]$.
Susceptibility factor	Analytical local susceptibility, clipped to $[0.1, 10]$; no normalization by the population mean.
Internal-change signal	$\tau_q = 10$ ms, $\tau_\pi = 80$ ms; within each trial, $\pi(t) = e^{-t/\tau_\pi}$.
Feedback pathway	$\tau_u = 15$ ms; $m_{\max} = 0.8$; $\alpha = 2$ in both training and testing; $\boldsymbol{m} = \alpha m_{\max} \tanh(\boldsymbol{u}/m_{\max})$; $\beta = 0.35$.
Associative input	Population-tuned context; $M_c = 32$, with $K_c = 2$ inputs per granule-cell unit; $\kappa_c = 4$, $\rho = 1.25$, fixed gain $\gamma_c = 0.937912$; window 100–300 ms.
Evaluation and randomness	Network seeds 41–45; states independently reset for each trial; P/W frozen and context withdrawn after training; pattern samples, input projections, and starting-point perturbations paired across conditions.

Table 3: Parameters of the baseline model, learning rules, and numerical protocol.

A.1 RECURRENT DRIVE AND FEEDBACK GAIN

Let p_J denote the recurrent connection density, g the connection gain, M_{ij} an independent Bernoulli connection mask, and ζ_{ij} a standard Gaussian variable. The fixed recurrent matrix is

$$J_{ij} = \frac{M_{ij} g \zeta_{ij}}{\sqrt{p_J N}}, \qquad M_{ij} \sim \text{Bernoulli}(p_J), \qquad \zeta_{ij} \sim \mathcal{N}(0, 1). \tag{10}$$

Let $[a]_+ = \max(a, 0)$. We decompose the current input contribution into two nonnegative quantities A_i^+ and A_i^-, and use the positive parameter β to specify the strength with which modulation changes

the effective gain:

$$A_i^+(\boldsymbol{r}) = \sum_j [J_{ij} r_j]_+, \qquad A_i^-(\boldsymbol{r}) = \sum_j [-J_{ij} r_j]_+,$$
$$F_i(\boldsymbol{r}, \boldsymbol{m}) = e^{\beta m_i} A_i^+(\boldsymbol{r}) - e^{-\beta m_i} A_i^-(\boldsymbol{r}). \tag{11}$$

The two terms summarize effective excitatory and inhibitory contributions at the current state, partitioned by the instantaneous direction of the input; they represent mixed populations at the pathway level. A positive m_i enhances the former and attenuates the latter, whereas a negative m_i produces the opposite modulation. When $\boldsymbol{m} = \mathbf{0}$, $\boldsymbol{F} = J\boldsymbol{r}$. The baseline uses $p_J = 0.1$, $g = 2$, and $\beta = 0.35$; recurrent connections remain fixed during training and evaluation.

A.2 Population coding in the auxiliary cortical population

The auxiliary population in Eq. (2) contains $M_c = 32$ units with preferred positions $\varphi_a = 2\pi a / M_c$. Class indices $\ell = 0, \ldots, 4$ correspond to a fixed random permutation of five equally spaced positions on the ring, denoted by θ_ℓ; the permutation separates class identity from ring order.

Each granule-cell unit selects $K_c = 2$ upstream inputs without replacement. The selected connections are $Q_{ka} = \omega_{ka}/K_c$, with $\omega_{ka} \sim U(0.8, 1.2)$, and all other connections are zero. The baseline uses tuning parameter $\kappa_c = 4$ and relative strength $\rho = 1.25$, corresponding to a predetermined fixed gain $\gamma_c = 0.9379119943$. The projection and gain remain fixed across trials. With $t_n = n\delta$ denoting a neural-step onset, steps satisfying $100 \le t_n < 300$ ms include associative input; associative input is zero during testing.

A.3 Constraints and numerical updates at the input site

The set $\mathcal{C}$ defined in Section 3.2 fixes the input topology, row sums, and individual edge bounds; Eq. (6) projects the weights onto this set after each accumulated update. The decay term reduces the displacement $P - P^0$, where P^0 is the uniform mapping at the start of training. The associative increment is given by Eq. (5); activity differences include changes within a neural step, and candidate increments are summed directly. For smooth activity, the associative component can be written as $\dot{P}_{kj}^{\text{assoc}} = \eta_P E_j^P \dot{h}_k$; jumps caused by binary Top-K switching are included through the finite differences in the actual update.

The presynaptic trace is set to zero at trial onset and propagated from $t = 0$ to the end of the P learning window. Postsynaptic activity is also recorded from trial onset, with the first observation serving as the differencing origin. Candidate increments are accumulated for $100 \le t_n < 300$ ms and submitted every $\Delta = 10$ ms according to Eq. (6). The actual submission times are $110, 120, \ldots, 300$ ms, for 20 submissions; the window includes activity changes after context onset, while the first activity difference after context withdrawal lies outside this window. P is updated after W learning at the same step, and the new weights participate in the dynamics from the next neural step.

A.4 Discrete internal-change sensing and its temporal kernel

Equation (7) accumulates changes in cortical activity at every neural step, with the leakage gate evaluated at the beginning of that step. The vector $\boldsymbol{q}$ is reset at each trial; the baseline uses $\tau_q = 10$ ms and $\tau_\pi = 80$ ms. For the continuous description, let $a(t) = \pi(t)/\tau_q$, with s and v as integration variables. The continuous expression in the main text gives, for zero initial condition,

$$\boldsymbol{q}(t) = \int_0^t \exp\left[-\int_s^t a(v)\, \mathrm{d}v\right] \dot{\boldsymbol{r}}(s)\, \mathrm{d}s. \tag{12}$$

This kernel retains both the direction and the time of activity changes. Leakage during the earlier phase makes transient changes decay rapidly; as the gate weakens, the preceding state displacement can persist to later times. Changes in opposite directions cancel in the integral, so the signal reflects the net history of changes rather than a simple sum of activity-change magnitudes.

A.5 Output plasticity and local feedback susceptibility

The output eligibility trace is propagated from trial onset over all G granule-cell units according to Eq. (8), with time constant $\tau_{E_W} = 400$ ms. At each submission, the $M_W = \min(G, 2K) = 3200$ rows with the largest $|\boldsymbol{E}^W|$ are selected; eligibility traces of the other rows remain stored. The baseline uses $\eta_W = 1.3 \times 10^{-7}\ \mathrm{ms}^{-1}$, $\lambda_W = 10^{-7}\ \mathrm{ms}^{-1}$, $\epsilon_E = 10^{-8}$, and $w_{\max} = 0.01$.

After the associative update in Eq. (9), decay and clipping are applied sequentially. The associative term directly uses the eligibility values of the selected rows; the numerical-validity threshold only determines whether decay is applied to a row, and all selected rows receive the final amplitude clipping. The learning window is 100–1000 ms, with actual submission times $110, 120, \ldots, 1000$ ms, for 90 submissions. Eligibility traces retain their native amplitude scale.

The susceptibility describes the local response of the cortical state to changes in the filtered cerebellar output at a fixed activity state. Let s_i^F denote the local change scale of the modulated recurrent drive, s_i^r the local slope of the cortical response, and s_i^u the local response gain of the downstream output. From Eqs. (11) and (4), take

$$\begin{aligned} s_i^F &= \beta \left[e^{\beta m_i} A_i^+(\boldsymbol{r}) + e^{-\beta m_i} A_i^-(\boldsymbol{r}) \right], \\ s_i^r &= 1 - r_i^2, \qquad s_i^u = |\alpha| \left[1 - \tanh^2(u_i / m_{\max}) \right]_+ . \end{aligned} \tag{13}$$

Writing $\widetilde{\chi}_i$ for the unbounded local susceptibility and clip for interval clipping, the one-step response scale and its bounded form are

$$\widetilde{\chi}_i = \frac{\delta}{\tau_r} |s_i^F s_i^r| s_i^u, \qquad \chi_i = \mathrm{clip}_{[0.1,10]}(\widetilde{\chi}_i). \tag{14}$$

The local effect of a given feedback change is generally larger when the adjustable recurrent drive is stronger and the cortical and downstream outputs are farther from saturation. This quantity is computed from the current state at each weight submission and preserves the amplitude scale of each unit. It multiplies q_i to modulate the change in W and does not participate in generating $\boldsymbol{q}$.

When one neural step advances from n to $n+1$, susceptibility uses the updated $\boldsymbol{r}_{n+1}$ and the $\boldsymbol{u}_{n+1}$ formed in that step, and the effective drive is evaluated at that cortical state. Equation (14) uses the filtered output state $\boldsymbol{u}$ as the local change variable and therefore does not include the single-step filtering coefficient from $W^\mathsf{T}\boldsymbol{h}$ to $\boldsymbol{u}$; output and eligibility histories are retained by their respective state equations. This is a one-step local response approximation; multi-step recurrent propagation is represented by subsequent closed-loop evolution.

B Training progress and interaction at the two sites

Class structure emerges progressively during joint learning (Fig. S1). To compare the continued contribution of the two sites at different stages, phase-specific freezing branches from the same checkpoint at trial 500 or 1000, and each branch continues training along the paired pattern and perturbation sequence. The within-seed silhouette difference between continued and frozen W learning decreases from 0.185 ± 0.107 at the trial-500 branch point to 0.034 ± 0.098 at trial 1000. When W is frozen late in training, the final silhouette coefficient is 0.437 ± 0.084 and the perturbation distance is 0.018 ± 0.002, indicating that the established output mapping can support subsequent learning at the input site.

To quantify nonadditivity in continued learning at the two sites, we branch from the common checkpoint at trial n and define

$$I_n = S_{PW} - S_{P_n,W} - S_{P,W_n} + S_{P_n,W_n}. \tag{15}$$

The subscript P_n or W_n indicates that the corresponding site is fixed after branching, while the other site continues training through trial 2000; the final term is the reference with both sites fixed. A positive value indicates that the silhouette gain from jointly continued learning exceeds the sum of the gains in the two single-site continuation branches. At branch points 500 and 1000, I_n is 0.191 ± 0.091 and 0.121 ± 0.109, respectively, and is positive for all five seeds. The nonadditivity of continued learning, together with the cross-mapping loss in the main text, supports a matching relationship between input representation and output feedback.

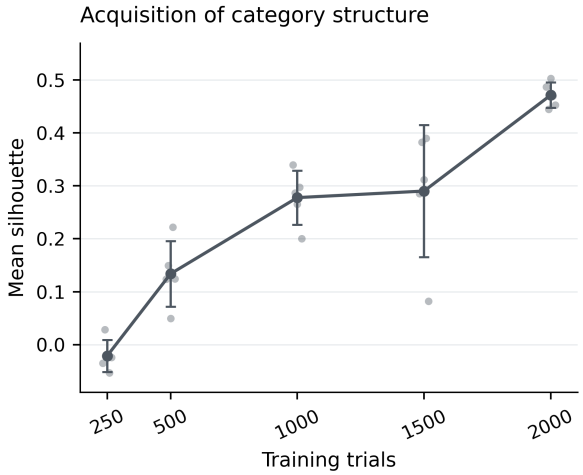


Figure S1: Emergence of class structure during joint learning. After trials 250, 500, 1000, 1500, and 2000, weights were frozen and context was withdrawn, and the five-class mean silhouette coefficient was measured. Points denote individual seeds; error bars show mean ±1 SD.

C GRANULE-CELL CODING AND CORTICAL DYNAMICAL CONDITIONS

Class organization and perturbation suppression depend non-monotonically on the granule-cell active fraction and cortical gain, and the two measures do not always improve together. To examine learning under different coding and dynamical conditions, we scanned Top-K active fractions from 1% to 15% in 2% increments and cortical gains $g = 0.75$ to 3 in increments of 0.25. The two scans fixed $g = 2$ and Top-$K = 5\%$, respectively, while all other learning and associative-input parameters were unchanged.

The mean silhouette coefficient reached the highest value in the tested range at Top-$K = 5\%$, 0.472 ± 0.024, and was 0.049 ± 0.052 and 0.168 ± 0.211 at 1% and 15%, respectively (Fig. S2**a**). The repeated-stimulus perturbation distance decreased from 0.393 ± 0.045 at 1% to 0.0242 ± 0.0145 at 5%, and further to 0.0066 ± 0.0042 at 15% (Fig. S2**b**). The cosine separation of granule-cell codes, S_{GC}, followed a similar trend to cortical class separation and was 0.471 ± 0.035 at 5% (Fig. S2**c**; definition in Appendix D.4). Better class organization at intermediate active fractions is consistent with the trade-off among dimensionality, feature sharing, and noise tolerance in expanded codes (Babadi & Sompolinsky, 2014; Litwin-Kumar et al., 2017; Cayco-Gajic & Silver, 2019). Because the amplitude of each binary code unit is fixed, the active fraction changes both total activity and eligibility-trace scale; these results therefore reflect joint changes in coding selectivity, learning, and feedback strength.

In the gain scan, the silhouette coefficient increased from -0.071 ± 0.015 at $g = 0.75$ to 0.472 ± 0.024 at $g = 2$, and then decreased to 0.265 ± 0.124 at $g = 3$ (Fig. S2**d**). At $g = 0.75$, the repeated-stimulus perturbation distance was close to zero (Fig. S2**e**), but class separation was negative, showing that convergence after a brief stimulus does not guarantee retention of class differences. The cumulative RMS of net trial-wise weight change, L_W, increased from 1.584×10^{-3} at $g = 2$ to 1.765×10^{-3} at $g = 3$, while class separation decreased (Fig. S2**f**). L_W measures the cumulative amplitude of net trial-wise changes and is not equivalent to the path length over all update steps or to learning speed. Across the tested conditions, larger weight changes or stronger repeated-stimulus consistency alone were insufficient to guarantee better class organization.

D EVALUATION METRICS AND STATISTICAL CALCULATIONS

Each metric is first calculated within an individual network seed, after which the mean and sample SD are reported across seed means. Classes are weighted equally; the treatment of samples and noise repetitions follows each metric definition below.

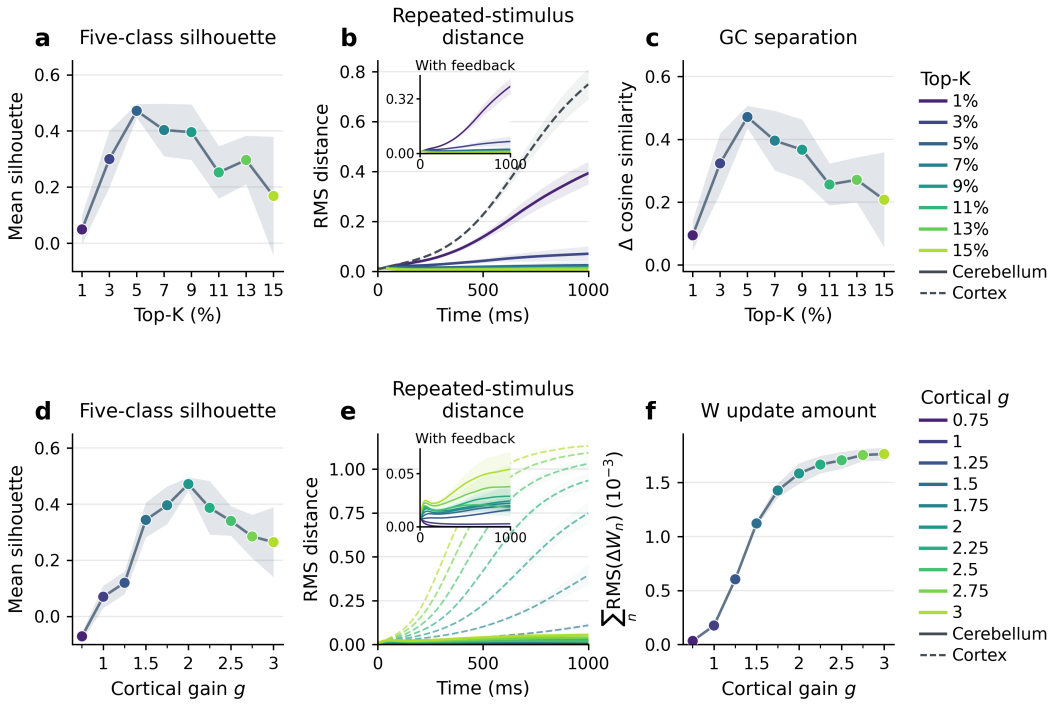


Figure S2: Granule-cell coding and cortical dynamical conditions. The top row scans Top-K with $g = 2$ fixed; the bottom row scans g with Top-$K = 5\%$ fixed. **a**, active fraction and silhouette coefficient. **b**, repeated-stimulus perturbation distance. **c**, cosine separation of granule-cell codes. **d**, gain and silhouette coefficient. **e**, corresponding perturbation distance. **f**, cumulative amplitude of net trial-wise W change. In **b** and **e**, solid lines show feedback after learning and dashed lines the corresponding cortex-only conditions; insets enlarge the feedback curves. Colors are identified by the legends at the right of each row.

D.1 Class silhouette coefficient and PCA visualization

Class organization is measured by the silhouette coefficient in the original cortical activity space. Let $\bar{\boldsymbol{r}}_i$ denote the mean activity of sample i at 1000 ms across its two repetitions, $d_{ij} = \|\bar{\boldsymbol{r}}_i - \bar{\boldsymbol{r}}_j\|_2$, and $\mathcal{I}_c$ the set of samples in class c. We define

$$a_i = \frac{\sum_{j \in \mathcal{I}_{c_i},\, j \neq i} d_{ij}}{|\mathcal{I}_{c_i}| - 1}, \qquad b_i = \min_{c \neq c_i} \frac{\sum_{j \in \mathcal{I}_c} d_{ij}}{|\mathcal{I}_c|}, \qquad s_i = \frac{b_i - a_i}{\max(a_i, b_i)}. \tag{16}$$

When the denominator is zero, $s_i = 0$. We first compute the within-class mean S_c and then use $S = \frac{1}{5}\sum_c S_c$ as the primary representation metric; the worst-class metric is $\min_c S_c$. A larger S indicates smaller within-class distances relative to between-class distances. PCA is used only for visualization, using a prespecified image set from a single seed, with 30 patterns per class and 150 final states from the first noise repetition; principal components are fitted independently for each condition, and quantitative silhouette coefficients are not computed in the reduced-dimensional space.

D.2 Trajectory speed, repeated distance, and internal-change magnitude

Trajectory speed, internal deviation, and repeated-stimulus perturbation distance are defined, respectively, as

$$v_i(t) = \mathrm{RMS}\left[\frac{\boldsymbol{r}_i(t) - \boldsymbol{r}_i(t-\delta)}{\delta}\right], \quad Q_i(t) = \mathrm{RMS}[\boldsymbol{q}_i(t)], \quad D_i(t) = \mathrm{RMS}[\boldsymbol{r}_i^{(1)}(t) - \boldsymbol{r}_i^{(2)}(t)], \tag{17}$$

where $\delta = 1$ ms, $\mathrm{RMS}(\boldsymbol{y}) = \sqrt{\sum_j y_j^2 / \dim(\boldsymbol{y})}$, and superscripts distinguish the two noise repetitions of the same sample. Speed is first calculated at each millisecond and then averaged within

each 10 ms recording interval; late-period speed is the mean over $800 < t \leq 1000$ ms. $D(1000)$ describes the separation between repetitions after one second. For learned and unlearned classes, we first average $Q_i(1000)$ across the two repetitions of the same sample to obtain a per-sample score, and then calculate $\mathrm{AUROC} = \Pr(Q_{\mathrm{OOD}} > Q_{\mathrm{ID}}) + \frac{1}{2}\Pr(Q_{\mathrm{OOD}} = Q_{\mathrm{ID}})$ for each seed.

D.3 Feedback drive and same-state drift

To distinguish feedback amplitude from its dynamical action, at the same state along the closed-loop trajectory we define $\boldsymbol{F}_0 = \boldsymbol{A}^+ - \boldsymbol{A}^-$ and $\boldsymbol{F}_m = e^{\beta \boldsymbol{m}} \odot \boldsymbol{A}^+ - e^{-\beta \boldsymbol{m}} \odot \boldsymbol{A}^-$, and calculate

$$R_{\mathrm{net}} = \frac{\mathrm{RMS}(\boldsymbol{F}_m - \boldsymbol{F}_0)}{\mathrm{RMS}(\boldsymbol{F}_0)}, \qquad R_{\mathrm{total}} = \frac{\mathrm{RMS}(\boldsymbol{F}_m - \boldsymbol{F}_0)}{\mathrm{RMS}(\boldsymbol{A}^+ + \boldsymbol{A}^-)}, \qquad R_{\mathrm{drift}} = \frac{\mathrm{RMS}(\boldsymbol{d}_m)}{\mathrm{RMS}(\boldsymbol{d}_0)}, \tag{18}$$

where $\boldsymbol{d}_m = (-\boldsymbol{x} + \boldsymbol{F}_m)/\tau_r$ and $\boldsymbol{d}_0 = (-\boldsymbol{x} + \boldsymbol{F}_0)/\tau_r$. These drifts describe $\dot{\boldsymbol{x}}$ and differ from $\dot{\boldsymbol{r}}$ in the speed metric; $\boldsymbol{d}_0$ is the counterfactual obtained by instantaneously removing modulation at the same closed-loop state, distinct from the autonomous trajectory obtained by disabling feedback from the starting point. Ratios are calculated per sample before averaging, with a numerical denominator floor of 10^{-12}.

Cancellation between effective excitatory and inhibitory contributions determines the scale of feedback changes relative to the net drive. Let $\boldsymbol{S}_{\mathrm{drive}} = \boldsymbol{A}^+ + \boldsymbol{A}^-$ and $\Delta\boldsymbol{F} = \boldsymbol{F}_m - \boldsymbol{F}_0$; componentwise,

$$\Delta\boldsymbol{F} = \sinh(\beta \boldsymbol{m}) \odot \boldsymbol{S}_{\mathrm{drive}} + [\cosh(\beta \boldsymbol{m}) - 1] \odot \boldsymbol{F}_0. \tag{19}$$

For small modulation, the first-order change is determined by the sum of effective excitatory and inhibitory drives, whereas the unmodulated net drive is determined by their difference. Thus, a change small relative to the total drive can still produce a large correction relative to the net drive. The same-state comparison in the main text retains the same cortical activity and feedback state to distinguish modulation amplitude from its immediate dynamical action.

D.4 Granule-cell code geometry and weight changes

Class selectivity of granule-cell codes is measured by cosine separation. Let C_{cc} denote the mean pairwise cosine similarity between final-time granule-cell codes $\boldsymbol{h}$ for different samples in class c, and C_{cd} the mean between classes c and d; then $S_{\mathrm{GC}} = \frac{1}{5}\sum_c [C_{cc} - \max_{d \neq c} C_{cd}]$. Both noise repetitions are retained, but within-class statistics exclude pairs corresponding to the same sample. Output-weight change is represented by $L_W = \sum_{n=1}^{2000} \mathrm{RMS}(W_n - W_{n-1})$, where W_n is the complete matrix after trial n; this measures the cumulative amplitude of net trial-wise changes and is not equivalent to the path length over all update steps or to task-learning speed.

D.5 Finite-amplitude perturbation propagation

To compare perturbation propagation near familiar and unlearned classes, we calculate the cumulative logarithmic gain of finite-amplitude perturbations in the frozen coupled network. Let the complete forward state be $s = (x, z, u)$, and use an RMS norm that weights the state blocks equally:

$$\|\delta s\|_* = \left[\frac{1}{3}\left(\frac{\|\delta x\|_2^2}{N} + \frac{\|\delta z\|_2^2}{G} + \frac{\|\delta u\|_2^2}{N}\right)\right]^{1/2}. \tag{20}$$

At $t_0 = 100\,\mathrm{ms}$, an independently sampled random direction is applied to each state block with $\|\delta s\|_* = \epsilon$; the distance between the two trajectories is then measured every $20\,\mathrm{ms}$, and the perturbation is rescaled to the same amplitude while preserving its evolved direction. Let d_j^- denote the distance before the jth rescaling and d_{j-1}^+ the measured distance after the preceding rescaling. Then

$$L_\epsilon(t_n) = \sum_{j=1}^{n} \log \frac{d_j^-}{d_{j-1}^+}. \tag{21}$$

$L_\epsilon < 0$ indicates cumulative contraction of the sampled perturbation directions under the specified norm; $\widehat{\lambda}_\epsilon = L_\epsilon/[(t - t_0)/1000]$ is the corresponding mean growth rate in s^{-1}. The mean local growth rate over the late interval 400–1000 ms is calculated as $[L_\epsilon(1000) - L_\epsilon(400)]/0.6$ to separate early transients. For each class, 10 test patterns are prespecified by their original sample indices,

with two random perturbation directions per pattern; results are first averaged over directions and within-class samples, and then aggregated with equal weights across classes and network seeds. The main analysis uses $\epsilon = 0.003$, and $\epsilon = 0.001$ is used to examine amplitude sensitivity. This metric follows the actual nonlinear trajectories and therefore retains winner switching in binary Top-K coding, describing perturbation propagation over finite times and amplitudes. Because the perturbation is repeatedly renormalized, e^{L_ϵ} is the cumulative product of local expansion or contraction factors, rather than the distance ratio of two trajectories that are never reset. Negative values over a finite window may also include contributions from rapid state decay and rotation of the perturbation direction, and are not equivalent to a negative maximal Lyapunov exponent or global asymptotic stability.

E Calibration and validation of learning strength

E.1 Structure of the recurrent replacement module

The cortical recurrent replacement module (RR) contains 1000 hidden units, recurrent connection density 0.5, and time constant 20 ms. Its input projection is dense; hidden activity passes through a $\tanh$ nonlinearity and produces feedback through a full output matrix, without Top-K competition. The actual spectral radius ρ_R of the internal recurrent matrix is 0.5, 1, 1.5, or 2, while the task-cortex recurrent gain is fixed at $g = 2$. RR and the cerebellar feedforward module (FF) use the same training patterns, trial order, starting-point perturbations, and associative-cue protocol, and are evaluated after context withdrawal. The comparison simultaneously changes expansion dimensionality, coding sparsity, and internal recurrence, thereby testing the joint performance of the two structural combinations.

For a computational-scale comparison, we estimated the effective single-step cost of the control pathway by counting multiply–accumulate (MAC) operations associated with nonzero connections, excluding the shared cortical recurrent update. With $G = 32{,}000$, $F = 16$, $K = 1{,}600$, and $N = H = 1{,}000$, FF requires $GF = 512{,}000$ MACs for the sparse P projection and $KN = 1{,}600{,}000$ MACs for the active W readout, giving approximately 2.11×10^6 MACs. Expansion-state updates and Top-K selection additionally involve elementwise, comparison, and indexing operations. RR requires $HN = 1{,}000{,}000$ MACs for the dense input projection, $0.5H^2 = 500{,}000$ for the recurrent matrix, and $HN = 1{,}000{,}000$ for the full readout, giving approximately 2.50×10^6 effective MACs. Elementwise nonlinearities and context additions are excluded from both counts. These estimates compare dominant arithmetic at the level of effective nonzero connections; they do not imply an exact FLOP, memory-traffic, or wall-clock match because sparse indexing and Top-K selection are implemented differently from dense matrix products.

E.2 Calibration of initial functional updates in feedforward and recurrent feedback modules

FF and RR use different hidden-activity codes, so identical learning rates do not ensure comparable functional changes from a single update. To match the initial update scale, we select the first 20 pattern samples from each class, 100 in total, and compute candidate updates under each of the five seeds. After each P or W submission, its effect is measured and the matrix is restored to its initial value before the next neural-dynamics step; thus, calibration does not accumulate weight learning and does not use test data.

Let a_k denote the activity variable used for calibration before the kth P submission, ΔP_k the candidate P update after imposing the constraints specific to each structure, and P_0 the initial input mapping. FF uses the baseline-shifted $a_k = r_k + 1$, whereas RR uses $a_k = r_k$. Because the FF update preserves row sums, $\Delta P_k \mathbf{1} = \mathbf{0}$, so the drive increment is unchanged by the shift; the normalization denominator is defined using the baseline-shifted activity. This convention specifies the calibration statistic separately from the granule-cell forward drive $P\boldsymbol{r}$. The initial functional change in P is

$$U_P = \left\langle \frac{\mathrm{RMS}(\Delta P_k a_k)}{\max\{\mathrm{RMS}(P_0 a_k), 10^{-12}\}} \right\rangle_k . \tag{22}$$

Here the matrix product for FF includes only its existing connections. For W, let h_k denote the readout activity at the state where the update occurs and ΔW_k the candidate W update; define

$$U_W = \left\langle \mathrm{RMS}(\Delta W_k^{\mathsf{T}} h_k) \right\rangle_k . \tag{23}$$

For FF, h_k is a binary Top-K code; for RR, it is the $\tanh$ hidden activity multiplied by the fixed gain $1/\sqrt{1000}$. This quantity measures the raw readout change before output filtering and saturation, and is not the final modulation applied to the cortex.

For each structure and seed, we measure 2000 P submissions and 9000 W submissions, first averaging within seed and then taking the mean across five seeds. The RR probe learning coefficients are both set to 10^{-4}. For each spectral radius, the two learning strengths are determined by

$$\eta_P^{\mathrm{RR}} = 10^{-4} \frac{\overline{U}_P^{\mathrm{FF}}}{\overline{U}_P^{\mathrm{RR,probe}}}, \qquad \eta_W^{\mathrm{RR}} = 10^{-4} \frac{\overline{U}_W^{\mathrm{FF}}}{\overline{U}_W^{\mathrm{RR,probe}}} \tag{24}$$

and the same calibration test is repeated with the resulting coefficients. The FF targets are $\overline{U}_P = 9.5955 \times 10^{-4}$ and $\overline{U}_W = 4.2406 \times 10^{-5}$. For all four RR spectral-radius conditions, the ratio of the calibrated five-seed mean to its target is within 10^{-6} of 1; matching individual seeds one by one is not a constraint of this calibration. Coefficients for each spectral radius are listed in Table 4.

RR actual spectral radius	P learning coefficient	W learning rate (ms^{-1})
0.5	3.55694×10^{-5}	4.22496×10^{-4}
1.0	3.22362×10^{-5}	3.45089×10^{-4}
1.5	2.74001×10^{-5}	2.74788×10^{-4}
2.0	2.23813×10^{-5}	2.37802×10^{-4}

Table 4: Calibration coefficients for the initial functional updates in RR. Each spectral radius uses learning strengths shared across seeds to match the mean effect of an initial single submission.

E.3 Learning-rate adjustment for the instantaneous-derivative teaching signal

The instantaneous-derivative comparison examines the temporal organization of the teaching signal by replacing q with $q^{\mathrm{deriv}}(t) = [r(t+\delta) - r(t)]/\delta$, while keeping the remainder of the W update rule unchanged. The original learning rate is $\eta_W^0 = 1.3 \times 10^{-7}\ \mathrm{ms}^{-1}$. Because the two signals differ in units and magnitude, learning-rate multipliers are estimated from the actual changes in weights and readouts.

Let W_n^- and W_n^+ denote the output matrices at the beginning and end of trial n, respectively, and let h_n^* be the fixed granule-cell code captured in that trial before context onset, at 100 ms. Define

$$\overline{\Delta}_W = \left\langle \mathrm{RMS}(W_n^+ - W_n^-) \right\rangle_n , \qquad \overline{\Delta}_u = \left\langle \mathrm{RMS}\big[(W_n^+ - W_n^-)^{\mathsf{T}} h_n^*\big] \right\rangle_n . \tag{25}$$

For each trial, h_n^* is held fixed when readouts before and after the update are computed, removing the direct contribution of P changes to the readout difference; reference codes across trials or training conditions are not constrained to have identical activity. $\overline{\Delta}_W$ is the net matrix change per trial rather than the cumulative path of changes submitted every 10 ms, and $\overline{\Delta}_u$ is the raw readout change without output filtering or saturation.

In the original-q and unscaled-derivative experiments, the five-seed means of $\overline{\Delta}_W$ are 7.9225×10^{-7} and 4.1012×10^{-9}, respectively, with a ratio of approximately 193; the corresponding values of $\overline{\Delta}_u$ are 2.4685×10^{-3} and 5.4819×10^{-6}, with a ratio of approximately 450. We therefore set

$$\eta_W^{(193)} = 193 \eta_W^0 = 2.509 \times 10^{-5}\ \mathrm{ms}^{-1}, \qquad \eta_W^{(450)} = 450 \eta_W^0 = 5.850 \times 10^{-5}\ \mathrm{ms}^{-1}. \tag{26}$$

Both multipliers are estimated from mean changes in the original experiments and held fixed in subsequent training. Each condition uses five seeds and the same 2000 training patterns, with pattern order, classes, and starting-point noise paired trial by trial. Each trial contains 20 P updates and 90 W updates, with context present from 100 to 300 ms; test patterns and the frozen-evaluation protocol are the same as in the main experiment.

Changing the learning rate changes the closed-loop dynamics and therefore the derivative and eligibility traces, so the actual changes need not scale linearly with the multiplier. Table 5 reports measurements after retraining: in the 193-fold condition, $\overline{\Delta}_W$ and $\overline{\Delta}_u$ are 61.6% and 55.3% of their respective values in the q condition; in the 450-fold condition, they are 99.0% and 93.6%, giving an approximate match when averaged over the full training sequence. The recorded fraction of clipped W weights is zero in all conditions, so the scale difference is not dominated by hard weight clipping.

Teaching condition	η_W (ms^{-1})	$\overline{\Delta}_W$ (10^{-7})	$\overline{\Delta}_u$ (10^{-3})
Accumulated deviation q	1.3×10^{-7}	7.922 ± 0.477	2.468 ± 0.134
Instantaneous derivative, original rate	1.3×10^{-7}	0.041 ± 0.001	0.00548 ± 0.00035
Instantaneous derivative, 193-fold	2.509×10^{-5}	4.878 ± 0.182	1.366 ± 0.059
Instantaneous derivative, 450-fold	5.850×10^{-5}	7.843 ± 0.310	2.310 ± 0.093

Table 5: Actual update scales after replacing the teaching signal. Values are first averaged across all training trials within each seed and then aggregated across networks; the table reports measurements after retraining with the listed learning rates.

Task performance at each learning strength is as follows. At the original learning rate, the instantaneous-derivative condition has a silhouette coefficient of -0.024 ± 0.010 and $D(1000) = 0.736 \pm 0.047$. With a 193-fold increase, $D(1000)$ decreases to 0.030 ± 0.011, with a silhouette coefficient of 0.187 ± 0.100; in the 450-fold condition, $D(1000)$ is $(7.32 \pm 7.68) \times 10^{-4}$ and the silhouette coefficient is -0.135 ± 0.083. As a reference, the q condition has a silhouette coefficient of 0.472 ± 0.024 and $D(1000) = 0.024 \pm 0.014$. After calibration, derivative-based teaching restores perturbation suppression, but none of the tested conditions simultaneously reaches the class separation obtained with q.

The matching above is an amplitude match averaged over the full training process and does not make the two teaching signals temporally equivalent. For example, in the 450-fold condition, $\overline{\Delta}_W$ over the first 100 trials is 1.803×10^{-6}, higher than 6.870×10^{-7} in the q condition; over the final 100 trials it is 4.299×10^{-7}, lower than 6.813×10^{-7} in the q condition. Thus, this experiment tests whether simply increasing the strength of derivative-based learning suffices to restore class organization, but cannot separate the contributions of teaching-signal direction, temporal allocation, and the resulting closed-loop activity differences.

F Reproducibility

Network initialization, input projections, and data splits are determined by independently recorded random seeds; learning windows, state initialization, and numerical discretization follow the definitions above. Parameter selection uses only training and validation data; an independent test set is used for final evaluation. Original-space metrics for different random seeds are calculated separately before aggregation, and PCA visualizations retain the coordinate system of a single seed.