Synaptic weight dynamics underlying systems consolidation of a memory

Brandon J Bhasin[a,b], Jennifer L Raymond[c,1], and Mark S Goldman[b,d,1]

[a]Department of Bioengineering, Stanford University, Stanford, CA 94305
[b]Center for Neuroscience, University of California, Davis, CA 95616
[c]Department of Neurobiology, Stanford University School of Medicine, Stanford, CA 94305
[d]Departments of Neurobiology, Physiology, and Behavior, and Ophthalmology and Vision Science, University of California, Davis, CA 95616
[1]Correspondence can be addressed to jenr@stanford.edu, msgoldman@ucdavis.edu

Abstract. Systems consolidation is a common feature of learning and memory systems, in which a long-term memory initially stored in one brain region becomes persistently stored in another region. We studied the dynamics of systems consolidation in simple circuit architectures modeling core features of many memory systems: an early- and late-learning brain region and two sites of plasticity. We show that the synaptic dynamics of the circuit during consolidation of an analog memory can be understood as a temporal integration process, by which transient changes in activity driven by plasticity in the early-learning area are accumulated into persistent synaptic changes at the late-learning site. This simple principle leads to two constraints on the circuit operation for consolidation to be implemented successfully. First, the plasticity rule at the late-learning site must stably support a continuum of possible outputs for a given input. We show that this is readily achieved by heterosynaptic but not standard Hebbian rules, that it naturally leads to a speed-accuracy tradeoff in systems consolidation, and that it provides a concrete circuit instantiation for how systems consolidation solves the stability-plasticity dilemma. Second, to turn off the consolidation process and prevent erroneous changes at the late-learning site, neural activity in the early-learning area must be reset to its baseline activity. We propose two biologically plausible implementations for this reset that suggest novel roles for core elements of the cerebellar circuit.

Significance Statement. How are memories transformed over time? We propose a simple organizing principle for how long term memories are moved from an initial to a final site of storage. We show that successful transfer occurs when the late site of memory storage is endowed with synaptic plasticity rules that stably accumulate changes in activity occurring at the early site of memory storage. We instantiate this principle in a simple computational model that is representative of brain circuits underlying a variety of behaviors. The model shows how neural circuits can store new memories while preserving core features of older ones, and suggests novel roles for core elements of the cerebellar circuit.

Introduction

Memory systems transform transiently present information into a more persistent form. In short-term (working) memory, transient spiking activity of input neurons is transformed into persistent activity in downstream short-term memory-storing circuits (Zylberberg & Strowbridge 2017; Goldman-Rakic 1995). In long-term memory, memories stored transiently through neural plasticity at one site may become stored persistently at another site, through a process known as systems consolidation (Dudai *et al.* 2015; Squire *et al.* 2015). Systems consolidation is a common feature

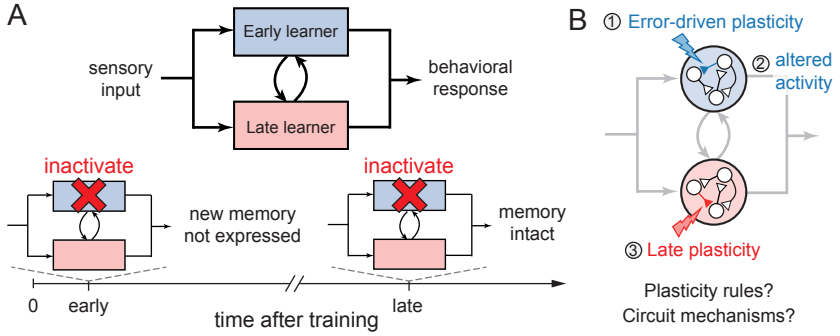

Figure 1: Schematic of systems consolidation. (A) Learned changes in the response to a sensory input occur first in an early-learning area, but are transferred over time to a late-learning area that becomes the sole site required for expression of learning, as revealed by inactivation experiments. (B) Feedback about behavioral errors drives synaptic plasticity in the early-learning area (1), leading to changes in activity (2) that in turn drive secondary plasticity in the late-learning area (3). Here, we investigate the dynamical mechanisms underlying this consolidation process and their implications for circuit organization and function.

of learning and memory systems, including declarative memory (Genzel & Wixted 2017), fear conditioning (Do Monte *et al.* 2016), and motor skill learning (Krakauer & Shadmehr 2006). It is thought to help memory systems navigate the "stability-plasticity" dilemma (Abraham & Robins 2005; Grossberg 1987)—balancing the need to have capacity for new memories and the tendency of new memories to overwrite old ones, which could lead to catastrophic forgetting (McClelland *et al.* 1995; Roxin & Fusi 2013).

In working memory, the transformation of transient representations into a more persistent form can be characterized by well-established computational principles (Brody *et al.* 2003; Chaudhuri & Fiete 2016; Goldman *et al.* 2009; Major & Tank 2004; Seung 1996; Wang 2001; Zylberberg & Strowbridge 2017). For the storage of analog (graded or continuous-valued) memories, this transformation can be accomplished through temporal integration of transient activity in the input to a circuit into persistent changes in circuit output, with the set of possible stored activity patterns forming a continuous set of stable patterns.

For systems consolidation, overarching computational principles describing the transformation of transient into persistent representations are less well established. Qualitatively, the standard view of systems consolidation suggests that transient plasticity in an early-learning brain area results in altered neural activity that then triggers the induction of persistent changes in the late-learning area post-training (Do Monte *et al.* 2016; Krakauer & Shadmehr 2006; Hardt *et al.* 2013; Lesburguères *et al.* 2011; Richards & Frankland 2017; Squire *et al.* 2015). Through this process, the expression of learning becomes robust to inactivation of the early-learning area (Fig. 1*A*,*B*).

We investigated the dynamics of systems consolidation in a model of a simple circuit that captures essential features of the systems consolidation of error-driven learning in brain areas such as the cerebellum (Cooke *et al.* 2004; Lisberger 2021; Raymond & Medina 2018), striatum (Andalman & Fee 2009; Warren *et al.* 2011; Yin *et al.* 2009; Makino *et al.* 2016; Teşileanu *et al.* 2017; Murray & Escola 2020), and amygdala (Medina *et al.* 2002; Do Monte *et al.* 2016). We frame systems

consolidation as a process of temporal integration, in which transient changes at the initial site of plasticity are integrated into changes in the final site. We show that this simple principle places strong constraints on learning that are met only by certain classes of plasticity rules, and provides insight into how a circuit may mitigate the stability-plasticity dilemma. Further, considered within the context of cerebellar systems consolidation, we show how the requirements for consolidation suggest novel roles for core elements of the cerebellar circuit.

Results

A simple circuit model of systems consolidation

We modeled a simple circuit that learns the analog-valued gain of an input-to-output transformation. This gain is determined jointly by a direct and an indirect pathway through the circuit. Errors in behavioral performance drive plasticity at an early-learning site in the indirect pathway, which is then consolidated at a late-learning site in the direct pathway.

To make the analysis more concrete, we consider the specific example of oculomotor learning. Our focus is not on the intricacies of oculomotor learning specifically but the principles of consolidation that recur in circuit architectures throughout the brain. Nevertheless, our model captures many important features of oculomotor learning. The gain of eye movement responses to vestibular or visual stimuli can be adaptively modified by learning so as to attain any value within an analog range (Fig. 2*A*; see Materials and methods, *Feedforward sensorimotor circuit model*; Broussard & Kassardjian 2004). Expression of this learned change in the sensory-to-motor transformation initially depends on the cerebellar cortex, but becomes cerebellum-independent within 24 hours post-training (Anzai *et al.* 2010; Jang *et al.* 2020; Kassardjian *et al.* 2005; Nagao & Kitazawa 2003; Shutoh *et al.* 2006). Learning and consolidation occur over timescales that are long compared to the eye movement responses to sensory stimuli, hence we modeled the latter as instantaneous.

We first studied a model of the circuit in which the input-to-output transformation was purely feedforward. In the context of cerebellum-dependent learning, this is consistent with the classical Marr-Albus-Ito model and previous computational studies of the consolidation of oculomotor learning (Albus 1971; Clopath *et al.* 2014; Herzfeld *et al.* 2020; Ito 1982; Marr 1969; Porrill & Dean 2007; Yamazaki *et al.* 2015). We simulated neural activity and changes in synaptic weights in the feedforward circuit model during a training and post-training period. During the training period, the model receives instructive signals about behavioral performance, which control the induction of plasticity at the early-learning site. For the case of oculomotor learning, these instructive signals reflect behavioral errors resulting from failures of eye movements to stabilize images on the retina, carried by the climbing fiber input to the cerebellar cortex. Such signals induce a decrease in the weight w_H at the early-learning site, reducing the inhibitory output from the indirect (cerebellar cortical) pathway and thereby increasing the overall gain of the sensory-to-motor transformation (Fig. 2*B,C*; Boyden *et al.* 2004; Inoshita & Hirano 2018; Jang *et al.* 2020; Kimpo *et al.* 2014).

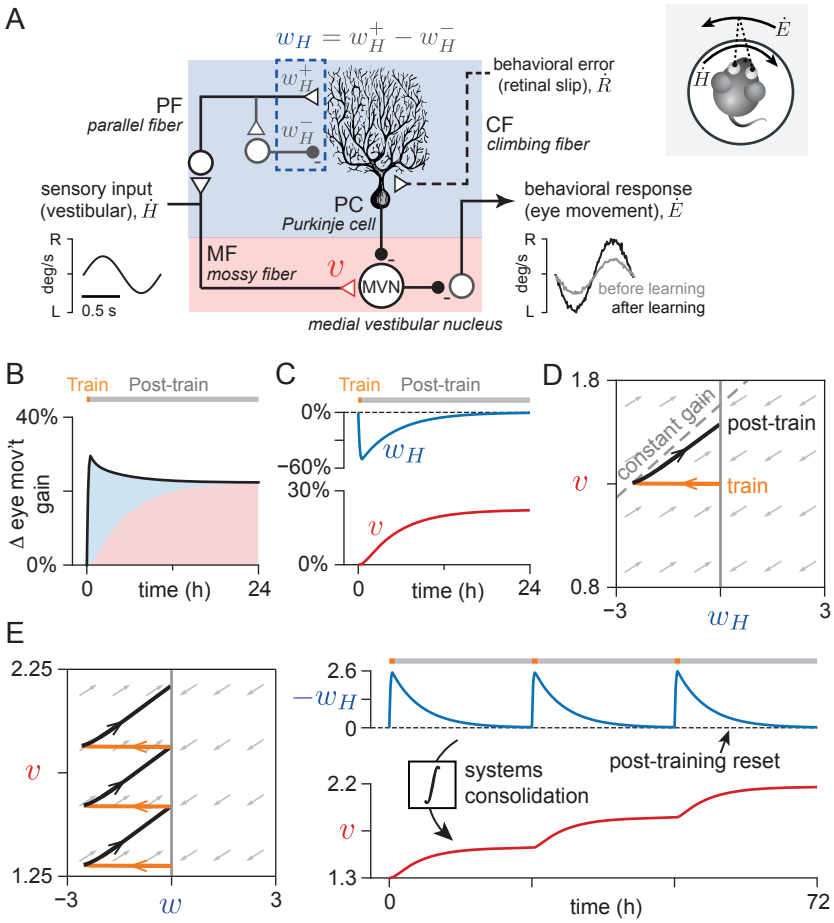

Figure 2: Synaptic weight dynamics in a representative feedforward-architecture circuit model of systems consolidation correspond to a temporal integration process. (*A*) The microcircuit underlying oculomotor learning alters the amplitude ("gain") of the reflexive eye movement response to sensory (vestibular) input. Incorrect eye movement amplitude causes image motion on the retina ("retinal slip"). Instructive signals carrying information about behavioral errors, i.e., retinal slip, drive changes in weight, w_H, at an early-learning site in the cerebellar cortex (blue shaded area), with w_H representing the difference between excitatory (w_H^+) and inhibitory (w_H^-) synaptic strengths. Over time, expression of learning becomes dependent only on the weight v at a late-learning site in the brainstem (red shaded area). (*B, C*) Simulation of model for 30 minutes of training (orange block) to increase the input-to-output gain of the eye movement response, followed by 23.5 hours post-training in the dark (grey block). During the post-training period, the model received no information about errors. (*B*) Change in the gain of the eye movement response (black line), with shading showing contributions from the early- (blue) and late-learning (red) sites. (*C*) Change in weights w_H (blue) and v (red). (*D*) Trajectory of synaptic weights during the training (orange) and post-training (black) periods. Grey arrows show the analytically calculated, approximate instantaneous direction in which a weight configuration at a given point will evolve without information about errors. All trajectories tend toward some marginally stable point along the line $w_H = 0$ (thick grey line). A trajectory resulting in perfect consolidation would follow the "constant gain" line (grey dashed line). (*E*) Consolidation corresponds to a temporal integration process in which transient changes in activity at the early-learning site, driven by plasticity at w_H, are accumulated into persistent changes in v. Panels show trajectory and dynamics in synaptic weight space (*left*) and over time (*right*) in a simulation of three consecutive days of the training protocol in *B–D*.

During the subsequent post-training period, information about behavioral errors is not available (experimental subject is placed in the dark to eliminate visual feedback about the stabilization of images on the retina), and we assume w_H decays back to a baseline value of zero (Fig. 2*C*). This baseline value of zero for w_H can be interpreted as a balance of excitation and inhibition at the early-learning site. The decay of plasticity at w_H is consistent with electrophysiological measurements (Jang *et al.* 2020), the experimental observation that memory expression becomes cerebellum-independent over time (Anzai *et al.* 2010; Kassardjian *et al.* 2005; McElligott *et al.* 1998; Nagao & Kitazawa 2003), and previous models (Clopath *et al.* 2014; Herzfeld *et al.* 2020; Yamazaki *et al.* 2015).

Because the synaptic changes induced at the early-learning site during the training period are transient, successful consolidation requires the induction of persistent changes at the late-learning site. Consolidation is known to depend on neural activity during the post-training period (Okamoto *et al.* 2011), which presumably induces the plasticity at the late-learning site (Jang *et al.* 2020). To achieve this, we implemented a heterosynaptic plasticity rule for the late-learning weight v in which weight changes are driven by the correlation of direct pathway input with early-learning area output (see Materials and methods, *Learning rules*), consistent with previous modeling of oculomotor learning (Clopath *et al.* 2014; Herzfeld *et al.* 2020; Medina & Mauk 1999; Menzies *et al.* 2010; Porrill & Dean 2007). As a result, altered post-training activity at the early-learning site induced an increase in v to a new steady-state value (Fig. 2*C*). This change persisted even as w_H returned to its baseline, supporting a persistent increase in the input-to-output gain of the circuit (Fig. 2*B,C*).

Dynamical principles of systems consolidation

To understand the features of the synaptic weight dynamics that support successful systems consolidation, we plotted the trajectory describing the joint evolution of the early- and late-learning weights. Initially, training induces a change at the early-learning site, decreasing the weight w_H from its baseline value and moving the weights (w_H, v) (Fig. 2*D*, orange trajectory) to a point corresponding to an increased input-to-output gain of the eye movement response (Fig. 2*D*, dashed "constant gain" line). During the post-training period, the early-learning weight decays to baseline, but the late-learning weight is driven to a new steady state value that preserves some of the increase in the input-to-output gain accrued during training (Fig. 2*D*, black trajectory). A nearly parallel trajectory will be followed by the evolution of the weights during the post-training period for any weight configuration reached during training (Fig. 2*D*, grey arrows), with all trajectories approaching some steady state value along a line in synaptic weight space (Fig. 2*D*, dark vertical line). Because the weight dynamics are at steady state whenever the activity of the early-learning area is at baseline, which occurs when the early-learning weight w_H decays to zero, there is a continuum of values that can be stably maintained by the late-learning weight v (see Materials and methods, *Analysis of synaptic learning dynamics*). Therefore, for a given initial value of v before

training, the final value reached after consolidation varies in a graded manner with the magnitude of the change in the early-learning weight w_H during training. This enables the circuit to learn and maintain any graded amplitude input-to-output gain across multiple learning events (Fig. 2*E*, S1; see Materials and methods, *Analysis of synaptic learning dynamics*).

These dynamics suggest an intuitive computational principle: systems consolidation of analog memories corresponds to a temporal integration process in which the late-learning weight stably accumulates changes induced during training at the early-learning site. For the simple feedforward model we have been examining, this can be visualized directly in the space of synaptic weights as an integration of transient changes in w_H into persistent changes in v (Fig. 2*E*). More generally, changes in the output of the early-learning area, driven by plasticity at w_H, are accumulated into weight changes at the late-learning site (see below for more complex circuit architectures).

To obey this principle, two conditions must be satisfied. First, the rule governing plasticity at the late-learning site must support the stable accumulation of persistent weight changes and corresponding continuum of input-to-output gains. Second, the circuit must reset the output of the early-learning site post-training so that the accumulation in the late-learning weight stops. Here we lay out the implications of the principle of consolidation as integration, starting with circuit function. Then, we show how the two simple conditions stated above place strong constraints on the plasticity rules and circuit features needed to support systems consolidation.

A speed-accuracy tradeoff for consolidation

Because the dynamics of systems consolidation can be understood as temporal integration, our results suggest that systems consolidation will exhibit properties that have been well-characterized in other kinds of integrators, such as the diffusive drift in the value being memorized due to the accumulation of noise, which has been found in neural integrator circuits implementing working memory (Brody *et al.* 2003; Burak & Fiete 2012; Ganguli *et al.* 2008; Lim & Goldman 2012; Seung 1996). In our model, noise in the neural activity driving plasticity at the late-learning weight v is accumulated, leading the input-to-output gain of the circuit to drift. In the absence of information about behavioral errors (e.g. during the post-training period), even if the output of the early-learning area is reset on average, noise in the early-learning weight w_H—for example, resulting from noise in the pathway that carries instructive signals during learning—is integrated into changes in v (Fig. 3*A,B*). The amplitude of changes in v post-training, resulting from changes in w_H, is defined by the slope of the flow field (Fig. 3*A,B*, grey arrows), which is proportional to the learning rate of v and inversely proportional to the rate at which the early-learning site w_H is reset to baseline. Therefore, reducing the learning rate at v reduces drift in v due to noise in w_H (Fig. 3*A–D*, S2*A*; Materials and methods, *Analysis of synaptic learning dynamics*). But reducing the learning rate at v also reduces the speed at which actual learning events are consolidated (Fig. 3*E,F*). Thus, more rapid consolidation of learning comes at the expense of a less accurate consolidated memory (Fig. 3*G*, S2*B*).

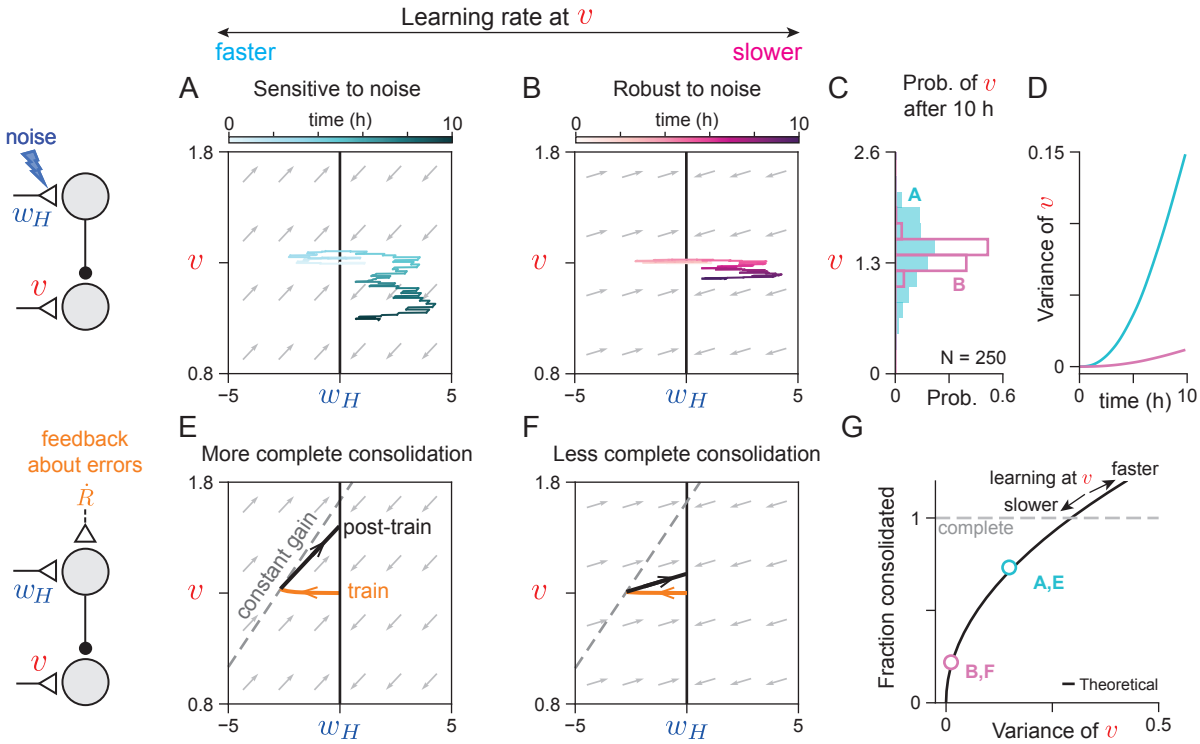

Figure 3: Diffusive drift of consolidated memory reflects a tradeoff between speed of consolidation and sensitivity to noise. (A, B) Trajectories of synaptic weights in response to periodic random perturbations in the early-learning weight w_H when the learning rate at v is relatively fast (A) or slow (B) (light to dark colors show progression of time). (C) Distribution of the value of the late-learning weight v at the end of 10 hours of random perturbations of w_H for $N = 250$ simulations (see Materials and methods for details), using the slower (open magenta bars) or faster (filled cyan bars) learning rates at v depicted in panels A and B, respectively. (D) Time course of the variance of the distributions in C. (E, F) Trajectories of synaptic weights during a training (orange) and subsequent post-training period (black). The learning rates in panels E and F were the same as in panels A and B, respectively. (G) Fraction of learning (change in gain during training that is consolidated post-training) compared to the variance of the weight v after 10 h of random perturbation. Circles show the empirical values of these quantities (taken from panels C, E and F) when the learning rate at v was relatively fast (cyan) or slow (magenta). Black curve shows the theoretical relationship between these quantities, as a function of the learning rate at v (see SI Text, S4, for derivation).

Consolidation mitigates the stability-plasticity dilemma through averaging

Framing consolidation as integration highlights mechanistically how systems consolidation can address the stability-plasticity dilemma (Abraham & Robins 2005; Grossberg 1987), i.e., the tradeoff between the ability of neural systems to adapt quickly and to not overwrite previous learning. Within the context of error-driven learning, this tradeoff arises when the optimal, "target" input-to-output gain of the circuit transformation, as conveyed by the instructive signals guiding learning, fluctuates across training sessions due to changes in the environment or in the motor plant (e.g., due to experimental manipulations, fatigue, injury, or changes in body mass). With one site of learning, a more plastic circuit that adapts quickly to reduce error during a given training session will tend to have a bigger initial error at the start of the next training session (Fig. 4*A–C*, "w_H fast"). On the other hand, a more stable circuit, whose gain adapts slowly and approximates the mean of the fluctuating learning target, will have, on average, smaller initial errors across training sessions, but will less fully reduce the error during a given training session (Fig. 4*A–C*, "w_H slow"; Materials and methods, *Mitigation of stability-plasticity dilemma with two sites of learning*).

With two sites of learning—a fast-learning and fast-forgetting early site of plasticity, and a slow-learning late site of consolidation—the circuit can mitigate this tradeoff. Slow consolidation at v stores the long-term average of the fluctuating target gain conveyed by the instructive signals across training sessions, while fast plasticity at w_H adapts quickly during a single training session but is reset post-training (Fig. 4*D*). In this way, the circuit can both minimize the average error at the start of each training session and respond quickly to reduce errors within a session (Fig. 4*E*; Materials and methods, *Mitigation of stability-plasticity dilemma with two sites of learning*). However, the tradeoff is not completely overcome; slow consolidation also means that the circuit responds slowly to changes in the mean of the varying learning target (Fig. 4*F*).

Implications for plasticity rules

To support consolidation, the late-learning site must be able to stably accumulate persistent weight changes over time. This requires that the plasticity rule at the late-learning site support a continuum of stable weight values, so that the synapse can persistently hold any change in the weight. This is readily achieved by a heterosynaptic plasticity rule (Fig. 2*E*), because weight changes stop whenever the activity at the early-learning site returns to its baseline or, more generally, becomes uncorrelated with the direct pathway input (Dean *et al.* 2002).

Recent work has alternatively proposed a Hebbian, covariance-like rule for the consolidation of oculomotor learning (Yamazaki *et al.* 2015), in which plasticity is proportional to presynaptic activity multiplied by the difference between postsynaptic activity and a sliding threshold equal to its recent average (see Materials and methods, *Learning rules*). Our analysis shows that this covariance rule can support a continuum of stable values, and hence consolidation, if the circuit receives no time-varying input or noise during the post-training period (Fig. 5*A,B*, light lines).

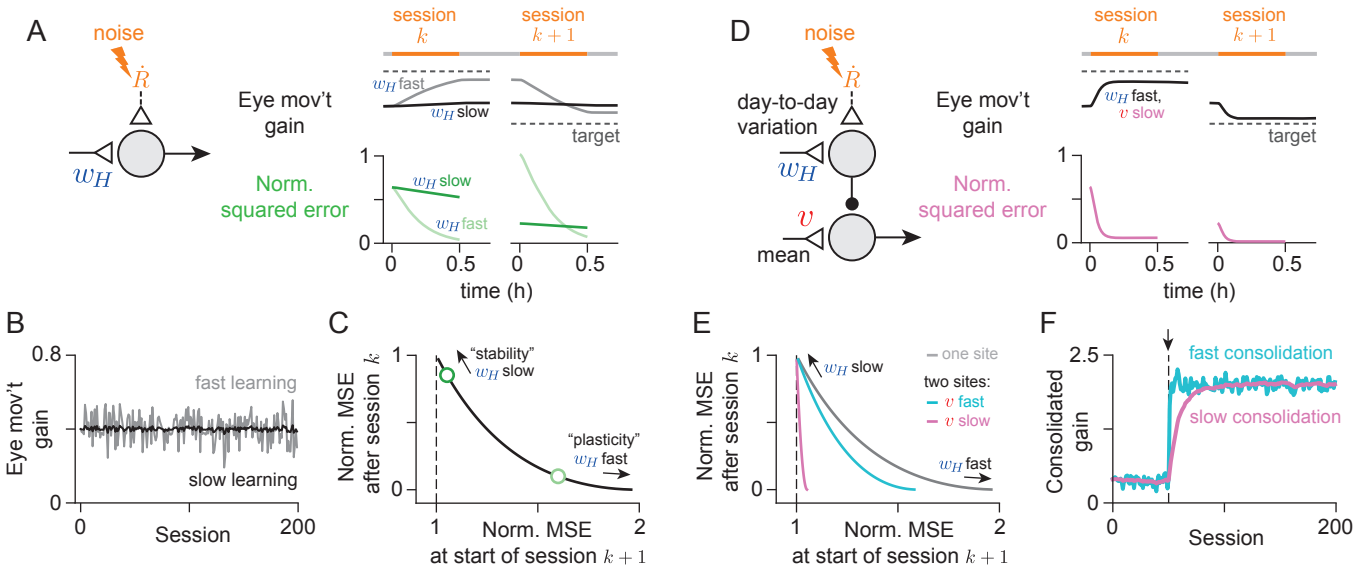

Figure 4: Consolidation averages over variability in the instructive signal to mitigate the stability-plasticity dilemma. (A) Learning in a model of a circuit with one site of plasticity that performs a graded amplitude input-to-output transformation is subject to the stability-plasticity dilemma. During a training session, the instructive signal drives a change in the weight w_H, which is stably remembered post-training. Across training sessions, the externally instructed target value of the input-to-output gain varies randomly about a mean value. Fast learning at w_H (light colors) reduces error more effectively during a training session compared to slow (dark colors), but increases the size of the expected error at the beginning of the next training session (distance from dashed line). (B) Input-to-output gain over 200 training sessions in a simplified model that simulates weight changes discretely over sessions (Materials and methods, *Mitigation of stability-plasticity dilemma with two sites of learning*). Slower learning (black) results in a more stable value of the weight due to greater averaging over variability in the instructive signal. (C) Expected (mean) squared error (MSE) after a training session in the simplified model trades off with expected error at start of subsequent session, calculated analytically and normalized to the variance of the target gain distribution. Dark and light green circles correspond to the learning rates used to generate black and grey curves in panel B. (D) A circuit with two sites of learning can mitigate the tradeoff in panel C. A persistent, slow-learning site of consolidation, v, can estimate the mean of the expected gain distribution, while a forgetful, fast-learning early site, w_H, can account for day-to-day variation in the instructive signal. (E) Relationship between normalized expected MSE at the end versus start of a training session when there are two sites of learning with fast (cyan) or slow (magenta) learning at the late site v. Grey curve replots the relationship when there is only one site of learning (same as C). (F) Consolidated gain in response to a change in the mean of the target gain distribution (at session indicated with arrow) for relatively fast (cyan) or slow (magenta) learning at v (rates same as D, but using the simplified model). Slow consolidation leads to reduced variability in the long run, but at the cost of larger errors early in training.

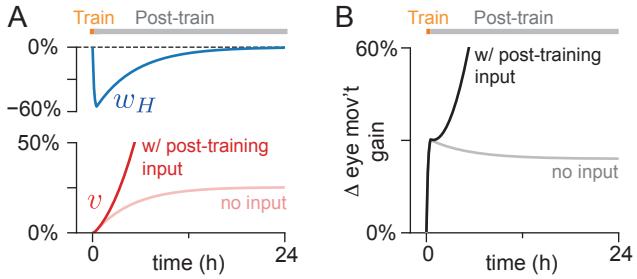

Figure 5: Hebbian covariance rule cannot support stable consolidation in the presence of time-varying signal or noise. (A) Evolution of the early- and late-learning weights, w_H (blue) and v (red), over time when using a Hebbian covariance rule at v, during a simulation of 0.5 h of training to increase the input-to-output gain of the circuit, followed by 23.5 h post-training without information about errors, when time-varying input to the circuit was either present (dark lines) or not present (light "no input" lines) post-training. (B) Change in the input-to-output gain corresponding to the simulations in A.

However, in the more biologically realistic case that the circuit does receive time-varying input or noise, the synaptic weight v, and thus the input-to-output gain, becomes unstable, growing exponentially during the post-training period (Fig. 5A,B, dark lines; SI Text, S1).

The failure of the covariance-like rule reflects an inherent source of instability in Hebbian learning (Miller & MacKay 1994). In the basic Hebbian rule, correlations between presynaptic and postsynaptic firing rates drive increases in synaptic weights. These increased synaptic weights in turn drive increased postsynaptic activity and thus increased correlations between pre- and postsynaptic firing, forming a positive feedback loop. For the covariance-like rule considered above, the sliding threshold counters the increased correlations associated with changes in the average postsynaptic firing rate, but does not counter the increased correlations associated with fluctuations around the average (Loewenstein 2008; see SI Text, S1). Other proposed methods for countering this instability, including weight normalization (Miller & MacKay 1994; Oja 1982) and firing rate-target homeostasis (Bienenstock *et al.* 1982; Chistiakova *et al.* 2015; Turrigiano 2008; Yger & Gilson 2015), also fail to support a continuum of stable values in our model. Rather, they yield dynamics that drive the late-learning weight to a single stable value (see SI Text, S2). Altogether, this analysis suggests that either heterosynaptic plasticity, or a different variant of Hebbian rule from those typically considered, is required to implement systems consolidation of an analog memory.

A circuit with an internal feedback loop requires an active post-training reset mechanism at the early-learning site

In addition to constraints on the form of plasticity at the late-learning site, consolidation also requires that the output of the early-learning area be reset during the post-training period, stopping the integration at the late-learning site. For the simple feedforward model described above, this condition is met by the intrinsic decay of the early-learning weight w_H to zero, which represents a balance of excitatory and inhibitory input to the early-learning area.

An active, rather than passive, post-training reset mechanism is necessary for stable consolidation in the more complex case where the input-to-output transformation computed by the circuit

contains an internal feedback loop (Fig. 6*A*). We modeled this case by adding to the model an internal feedback connection from the late- to the early-learning area. For oculomotor learning, the internal loop corresponds to efference copy of the eye movement command (Miles & Lisberger 1981; Lisberger & Sejnowski 1992; Lisberger 1994; Person 2019). Here we describe a model with a fixed internal feedback weight $w_E > 0$; similar results are obtained if w_E is plastic (Materials and methods, *Plasticity of internal feedback weight*; Fig. S3; SI Text, S6). As in the feedforward model, the condition that plasticity at the late-learning site support stable accumulation can be met by a heterosynaptic plasticity rule at v. However, the reset condition can no longer be met just by the passive decay of the weight w_H of the feedforward pathway in the early-learning area to a fixed baseline, but rather requires that, during the post-training period, w_H returns to a value that depends on v. This is because, due to the internal feedback from the late- to the early-learning area, changes in v also drive altered activity of the early-learning area. These changes must be offset post-training by changes in w_H so the feedforward sensory input and the feedback from the late-learning area effectively cancel, eliminating the drive for plasticity at the late-learning site and enabling v to reach a new steady state.

The reset requirement suggests a role for specific features of the cerebellar circuit architecture in consolidation

Active reset of the early-learning site could be achieved in at least two ways. First, the feedforward input to the early-learning area can be decomposed into the sum of a direct excitatory and disynaptic inhibitory pathway that have weights w_H^+ and w_H^-, respectively, controlled by separate plasticity mechanisms (Fig. 6*A*). We assume that changes in the excitatory weight are driven only by instructive signals occurring during the training period, with no passive decay. We then find that, during the post-training period, plasticity in the inhibitory weight can reset the activity of the early learning area such that it no longer responds to sensory input. This resetting is achieved when the inhibitory input weight is governed by a Hebbian rule driven by correlations between presynaptic inhibitory inputs and postsynaptic spiking relative to a sliding threshold (Fig. 6*B*–*E*; Materials and methods, *Reset Mechanism 1: Plasticity of inhibitory input to early-learning area*; Vogels *et al.* 2011). Alternatively, the reset could be driven by inhibition from the late-learning area onto the source of the instructive signals that control plasticity at the early-learning site (Fig. 6*F*–*J*; Materials and methods, *Reset Mechanism 2: Inhibition of pathway carrying instructive signals*). In the cerebellar context, the first reset mechanism would correspond to plasticity of molecular layer interneuron-to-Purkinje cell synapses, and the second could be achieved by the inhibitory pathway from the cerebellar nuclei to the inferior olive (see Discussion). Thus, a central computational role for these components of the cerebellar circuitry may be to implement the post-training reset of the rapid, early plasticity in the cerebellar cortex to support stable consolidation of learning.

Our model predicts that, following consolidation of learning to increase the input-to-output gain of the circuit, the sensitivity of cells in the early-learning area to feedforward input, w_H, will be

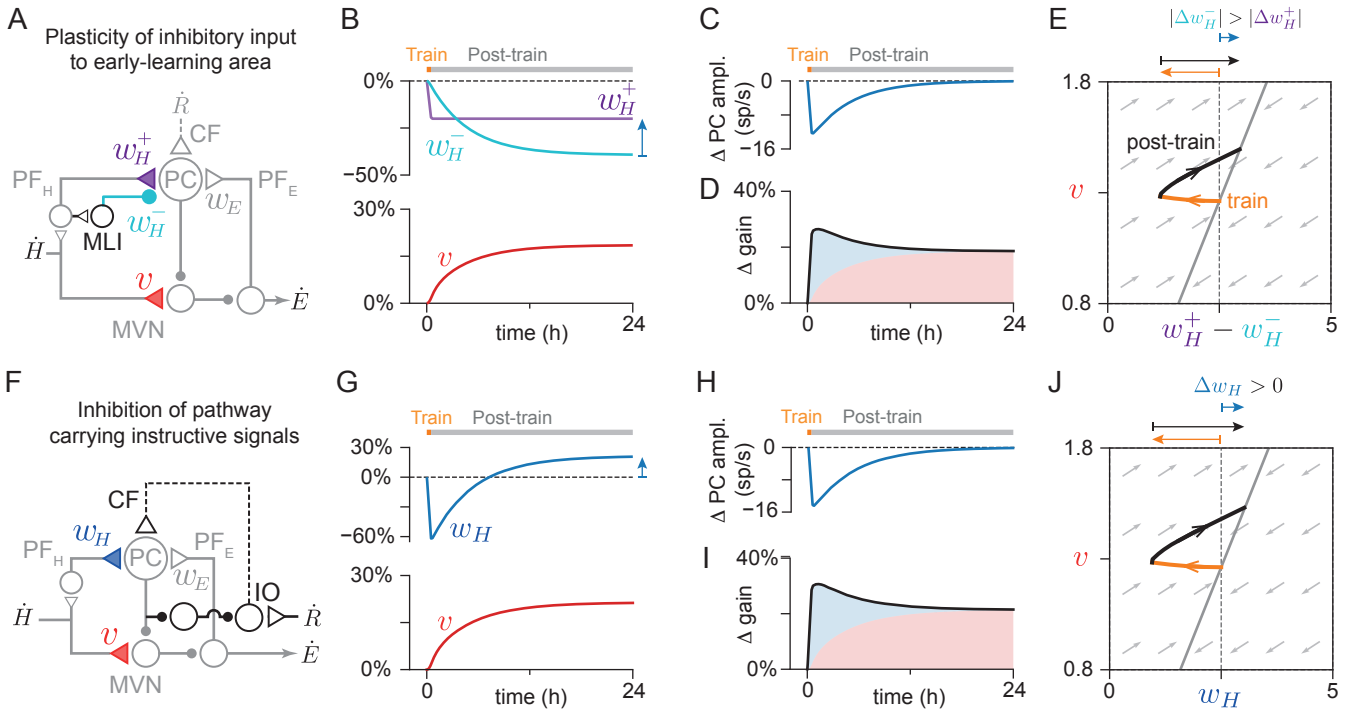

Figure 6: Circuit mechanisms for post-training reset enable consolidation in the presence of an internal feedback loop. (A–E) Reset of early-learning area output via plasticity of inhibitory input onto the early-learning area. (A) Circuit diagram. PF_H and PF_E: cerebellar parallel fiber input to Purkinje cells (PC) carrying sensory (vestibular) input $\dot{H}$ and efference copy feedback of motor output $\dot{E}$. MLI: molecular layer interneuron. CF: climbing fiber. $\dot{R}$: information about behavioral (retinal slip) errors (B) Change in early-learning excitatory weight w_H^+ (purple), inhibitory weight w_H^- (cyan) and late-learning weight v (red), during 0.5 h of simulated training (orange block) to increase the input-to-output gain of the circuit, followed by 23.5 h post-training with no information about errors (grey block). At steady state, the value of w_H^- has decreased more than w_H^+, so that the net change in the net weight $w_H = w_H^+ - w_H^-$ is positive (blue arrow). (C) Early-learning area output (amplitude of Purkinje cell activity relative to moving baseline), which drives consolidation at the late-site. (D) Change in the gain of the eye movement response (black line) to a sensory input. Blue and red shaded areas show the contribution of the early- and late-learning areas to the circuit transformation. (E) Trajectory of synaptic weights during the training (orange) and post-training (black) periods. During training, the net early-learning weight decreases (orange arrow above plot) from its initial value (dashed black line), but post-training (black arrow) approaches a steady state value that is larger than before training (blue arrow). Grey arrows show the approximate instantaneous direction in which a weight configuration at a given point in synaptic weight space will evolve during the post-training period, determined analytically. All trajectories tend toward a marginally stable point along a line (solid grey line). (F–J) Reset of early-learning area output via inhibition of the pathway carrying instructive signals to the early-learning site. (F) Circuit diagram. IO: inferior olive. (G) Change in early-learning weight w_H (dark blue) and late-learning weight v (red). The post-training reset drives w_H to a steady state value larger than before training (blue arrow). (H–J) Same as C–E, but for model shown in F.

higher than the pre-training baseline, as has been suggested based on experimental measurements (Miles *et al.* 1980; Lisberger 1994). This can be understood more clearly by plotting the dynamics of the synaptic weights. During consolidation, the post-training reset dynamics cause the weights to move along a trajectory to a new steady state along a line in the space of synaptic weights. In the feedforward model, this line is vertical, so the early-learning weight returns to its pre-training baseline during the post-training consolidation period (Fig. 2*E*). When there is an internal feedback loop, the line has a finite, positive slope, so the net weight w_H does not just return to its pre-training baseline value, but goes to a steady state value that is larger than before training (Fig. 6*E,J*, black arrows). Thus, our model predicts that, after consolidation, the change in the early-learning weight, relative to pre-training, will be in the opposite direction from the change in weight during training (Fig. 6*E,J*, blue vs orange arrows above plot; Payne *et al.* (2024)).

Discussion

We propose a computational principle governing systems consolidation of analog memories: systems consolidation is defined by a temporal integration process in which the late-learning weight stably accumulates changes induced during training at the early-learning site. To obey this principle, the circuit must have certain properties: First, the synaptic weight at the late-learning site must be governed by plasticity rules that enable it to stably accumulate and maintain any one of a continuum of values. We show that this is achieved by a heterosynaptic plasticity rule but not by standard stabilized Hebbian rules. Second, the output of the early-learning area must be reset post-training to stop the accumulation of weight changes at the late-learning site.

The accumulation of synaptic weight changes underlying systems consolidation is reminiscent of the temporal integration of transient spiking activity into persistent spiking activity in neural circuits that accumulate and store information in working memory, such as that implicated in the accumulation and storage of neural signals encoding evidence in decision-making tasks (Brody *et al.* 2003; Chaudhuri & Fiete 2016; Churchland & Ditterich 2012; Gold & Shadlen 2007; Goldman *et al.* 2009; Major & Tank 2004; Usher & McClelland 2001; Wang 2001; Wang 2008; Zylberberg & Strowbridge 2017). Our model shares distinctive features with such neural integrator circuits, arising from their similar underlying dynamics. First, stability of the memory requires that input to the integrator be precisely reset to zero post-training. Neural integrator circuits will integrate tonic background spiking activity unless these inputs are subtracted out or otherwise nullified (Cannon *et al.* 1983; Goldman *et al.* 2009; Seung 1996). In the case of systems consolidation, output from the early-learning area must be reset post-training; otherwise, it will drive incessant weight changes at the late-learning site. This is an example of a fine-tuning condition, a well-characterized property of integrators (Chaudhuri & Fiete 2016; Goldman *et al.* 2009; Seung 1996; Zylberberg & Strowbridge 2017). Drawing on the neural integrator literature, one proposed solution to this fine-tuning problem is to turn the continuum of steady states into a set of (many) discrete stable

fixed points (Goldman *et al.* 2003; Koulakov *et al.* 2002). This could correspond to our single "lumped" synaptic weight v being composed of the sum of the binary weights of individual bistable synapses (O'Connor *et al.* 2005) that are recruited with different thresholds (Goldman *et al.* 2003; Nikitchenko & Koulakov 2008). Second, noisy input causes diffusive drift that progressively corrupts the stored memory (Fig. 3; Burak & Fiete 2012; Ganguli *et al.* 2008; Lim & Goldman 2012). In our model, this led to a tradeoff between the speed of consolidation and the accuracy of the consolidated memory (Fig. 3*G*; Materials and methods, *Analysis of synaptic learning dynamics*). This can be interpreted as the long-term memory analog of the speed-accuracy tradeoff in decision-making tasks (Bogacz *et al.* 2010; Gold & Shadlen 2007).

The similarities in dynamics between neural integrator circuits and our systems consolidation model arise from the fact that in both cases, the memory stored in the circuit is analog—in our case, the amplitude of the circuit's learned response to a given input is graded. Much of the previous theoretical work on systems consolidation has instead considered the case of effectively binary neurons (Alvarez & Squire 1994; Murray & Escola 2020; Remme *et al.* 2021; Roxin & Fusi 2013; Tomé *et al.* 2022; Wittenberg *et al.* 2002), in which memory expression is evaluated with respect to whether or not a given neuron fires, rather than the graded value of its firing rate. In such models, Hebbian plasticity typically drives synaptic weights to take either a high (saturated) or a low value, making individual synaptic dynamics effectively bistable (Dong & Hopfield 1992; Miller & MacKay 1994). By contrast, in our model, the synaptic weight at the late-learning site must be able to stably take any analog value within a continuum, which is naturally achieved by a heterosynaptic plasticity rule. Although such heterosynaptic rules are commonly used in modeling supervised learning tasks (Dayan & Abbott 2005), they may form a more specialized class of biological plasticity than Hebbian learning rules. Supporting the continuum of steady states required for consolidation is more challenging for classical Hebbian rules (see SI Text, S2), though may be possible with more complex forms of Hebbian plasticity, such as a three-factor (Kuśmierz *et al.* 2017) or dendritic plasticity rule (Urbanczik & Senn 2014). Alternatively, as noted in the previous paragraph, it is possible that the apparent continuum of weight values modeled here instead reflects a sum of discrete, bistable synaptic contributions.

The framing of systems consolidation as a process of temporal integration provides mechanistic insight into previous work on how consolidation addresses the stability-plasticity dilemma (Abraham & Robins 2005; Grossberg 1987). In our model, the stability-plasticity dilemma is mitigated because, in the presence of variability in the instructive signals about behavioral errors, slow integration (i.e., averaging) at the late-learning site tracks long-timescale behavioral requirements, while the early-learning site quickly tracks short-timescale changes in these requirements (Fig. 4*D*; for a related use of integration to improve deep network training, see Johnson & Zhang 2013). In this manner, averaging leads to a more general memory at the late-learning site that is constantly updated by specific but transient memories at the early-learning site (Sekeres *et al.* 2018; Tse *et al.* 2007; Lindsey & Litwin-Kumar 2023; Richards & Frankland 2017; Sun *et al.* 2023). This framing

could also be extended to describe synaptic consolidation, through which transient early plasticity is consolidated into persistent, graded late plasticity in the same synapse (Jacquerie *et al.* 2023; Leimer *et al.* 2019; Li & Van Rossum 2020). The timescale over which the late-learning site averages is defined by the relative rates of plasticity at the late- and early-learning sites (Fig. 4*F*). This may suggest that the rate at which the circuit consolidates (i.e., the fraction of learning consolidated post-training) is tuned so that the timescale of the average is matched to the timescale over which the mean is expected to change in the world (Körding *et al.* 2007).

Though systems consolidation appears to be a common feature of learning and memory systems, the details of how it occurs in each system may be shaped by specific computational needs. In hippocampus-dependent memory, the early-learning circuit of the hippocampus learns the associations between components or features of an episode. Post-training replay of the activity patterns representing these associations then drives consolidation to the neocortex (Carr *et al.* 2011; Squire *et al.* 2015; Alvarez & Squire 1994; Tomé *et al.* 2022; Roxin & Fusi 2013; Murre 1996; Wittenberg *et al.* 2002). In the system we model, the circuit learns associations between a sensory input and a behavioral error signal that enforces a desired output. Post-training "replay" of the activity patterns learned at the early site then drives consolidation of the learned input-to-output transformation at the late-learning site. Although in both cases the correlations present during training are recapitulated in the post-training neural activity, hippocampal replay occurs as discrete events with transient, spontaneous reactivation of representations (e.g., during a sharp wave ripple), whereas in the cerebellum the "replay" could be driven by ongoing input to the circuit.

Our work extends previous modeling of the consolidation of oculomotor learning (Clopath *et al.* 2014; Herzfeld *et al.* 2020; Medina & Mauk 1999; Porrill & Dean 2007; Yamazaki *et al.* 2015) in two key ways. First, we analyzed the conditions under which consolidation can occur successfully. We showed that consolidation requires that plasticity at the late-learning site support a continuum of steady state weight values, a requirement readily met by a heterosynaptic plasticity rule. This explains why heterosynaptic rules have been successful in previous work (Clopath *et al.* 2014; Herzfeld *et al.* 2020; Medina & Mauk 1999; Porrill & Dean 2007), and is consistent with experimental work suggesting that plasticity at the late-learning site for cerebellum-dependent learning may be heterosynaptic (Menzies *et al.* 2010; Pugh & Raman 2006). Second, we considered an internal feedback loop from the late- to the early-learning area of the circuit, in contrast to the feedforward architecture used in previous models. This internal feedback loop, which carries efference copy of motor commands, has been suggested to be essential for producing the dynamics of individual eye movement responses (Lisberger & Sejnowski 1992; Lisberger 1994). Our model shows how, even in the presence of signals carried by this internal feedback, the output of the early-learning area can be reset post-training. We propose two biologically plausible circuit mechanisms for this reset, suggesting new functional roles for elements of the cerebellar microcircuit. The first mechanism, plasticity of the inhibitory feedforward inputs to the early-learning area, could be implemented by a Hebbian rule at the molecular layer interneuron (MLI)-to-Purkinje cell synapses (Vogels *et al.*

2011). Although these synapses are known to be plastic (Hirano 2018; Mapelli *et al.* 2015), weight changes were believed to be governed by climbing fiber-driven complex spikes in the postsynaptic Purkinje cell. Our model predicts that weight changes may instead be driven by the correlation of presynaptic MLI activity with Purkinje cell simple spikes. The second mechanism, inhibition of instructive signals by the late-learning area, could be implemented by an inhibitory pathway from the vestibular and cerebellar nuclei targeted by Purkinje cells to the inferior olive (Herzfeld *et al.* 2020; Najac & Raman 2015; Uusisaari & Knöpfel 2011), and might explain nonvisual climbing fiber responses that have been observed in the oculomotor cerebellum (Fanning *et al.* 2021; Winkelman & Frens 2006).

Learning and memory systems enable an organism to transform transient experiences into persistent effects. In working memory, neural integrator circuits accumulate and store transient input signals as persistent spiking activity. Here, we extend the concept of neural integration to the context of long-term memory by framing systems consolidation as a temporal integration process. Thus, this work suggests a unifying conceptual framework for describing computations underlying short- and long-term memory function.

Acknowledgements

We thank Alireza Alemi, Hannah Payne, Friedemann Zenke, Daniel Fisher, Emre Aksay, Brian Wiltgen, Michale Fee, Ilana Witten, David Tank and Jay McClelland for helpful insight and comments. This work was supported by a Stanford Interdisciplinary Graduate Fellowship and a Stanford Center for Mind, Brain, Computation and Technology Traineeship to BJB, NIH grant R01 DC004154-16S1 to JLR, a esearch to Prevent Blindness grant to UC Davis Ophthalmology and NIH grant R01 EY021581 to MSG, and Simons Collaboration on the Global Brain grant 543031 and NIH grants R01 EY031972 and R01 NS072406 to JLR and MSG. We assessed the gender and racial representation of authors we cite following the approach outlined in Dworkin *et al.* (2020), using open source code to evaluate the first names of the first and last authors (Zhou *et al.* 2022). Excluding self-citations to the first and last authors of our current paper, our references in the main text contain: 7.62% woman (first author)/woman (last author), 8.57% man/woman, 14.29% woman/man, and 69.52% man/man; and 12.99% author of color /author of color, 15.97% white author/author of color, 17.79% author of color/white author, and 53.26% white author/white author.

References

1. Zylberberg, J. & Strowbridge, B. W. Mechanisms of Persistent Activity in Cortical Circuits: Possible Neural Substrates for Working Memory. *Annual Review of Neuroscience* **40,** 603–627. doi:10.1146/annurev-neuro-070815-014006 (2017).
2. Goldman-Rakic, P. Cellular Basis of Working Memory. *Neuron* **14,** 477–485. doi:10.1016/0896-6273(95)90304-6 (1995).

3. Dudai, Y., Karni, A. & Born, J. The Consolidation and Transformation of Memory. *Neuron* **88,** 20–32. doi:10.1016/j.neuron.2015.09.004 (2015).

4. Squire, L. R., Genzel, L., Wixted, J. T. & Morris, R. G. Memory Consolidation. *Cold Spring Harbor Perspectives in Biology* **7,** a021766. doi:10.1101/cshperspect.a021766 (2015).

5. Genzel, L. & Wixted, J. T. in *Cognitive Neuroscience of Memory Consolidation* (eds Axmacher, N. & Rasch, B.) 3–16 (Springer International Publishing, Cham, 2017). doi:10.1007/978-3-319-45066-7_1.

6. Do Monte, F. H., Quirk, G. J., Li, B. & Penzo, M. A. Retrieving Fear Memories, as Time Goes By... *Molecular Psychiatry* **21,** 1027–1036. doi:10.1038/mp.2016.78 (2016).

7. Krakauer, J. W. & Shadmehr, R. Consolidation of Motor Memory. *Trends in Neurosciences* **29,** 58–64. doi:10.1016/j.tins.2005.10.003 (2006).

8. Abraham, W. C. & Robins, A. Memory Retention – the Synaptic Stability versus Plasticity Dilemma. *Trends in Neurosciences* **28,** 73–78. doi:10.1016/j.tins.2004.12.003 (2005).

9. Grossberg, S. Competitive Learning: From Interactive Activation to Adaptive Resonance. *Cognitive Science* **11,** 23–63. doi:10.1111/j.1551-6708.1987.tb00862.x (1987).

10. McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. Why There Are Complementary Learning Systems in the Hippocampus and Neocortex: Insights from the Successes and Failures of Connectionist Models of Learning and Memory. *Psychological Review* **102,** 419–457. doi:10.1037/0033-295X.102.3.419 (1995).

11. Roxin, A. & Fusi, S. Efficient Partitioning of Memory Systems and Its Importance for Memory Consolidation. *PLoS Computational Biology* **9,** e1003146. doi:10.1371/journal.pcbi.1003146 (2013).

12. Brody, C. D., Romo, R. & Kepecs, A. Basic Mechanisms for Graded Persistent Activity: Discrete Attractors, Continuous Attractors, and Dynamic Representations. *Current Opinion in Neurobiology* **13,** 204–211. doi:10.1016/S0959-4388(03)00050-3 (2003).

13. Chaudhuri, R. & Fiete, I. Computational Principles of Memory. *Nature Neuroscience* **19,** 394–403. doi:10.1038/nn.4237 (2016).

14. Goldman, M. S., Compte, A. & Wang, X.-J. in *Encyclopedia of Neuroscience* 165–178 (Elsevier, 2009). doi:10.1016/B978-008045046-9.01434-0.

15. Major, G. & Tank, D. Persistent Neural Activity: Prevalence and Mechanisms. *Current Opinion in Neurobiology* **14,** 675–684. doi:10.1016/j.conb.2004.10.017 (2004).

16. Seung, H. S. How the Brain Keeps the Eyes Still. *Proceedings of the National Academy of Sciences* **93,** 13339–13344. doi:10.1073/pnas.93.23.13339 (1996).

17. Wang, X.-J. Synaptic Reverberation Underlying Mnemonic Persistent Activity. *Trends in Neurosciences* **24,** 455–463. doi:10.1016/S0166-2236(00)01868-3 (2001).

18. Hardt, O., Nader, K. & Nadel, L. Decay Happens: The Role of Active Forgetting in Memory. *Trends in Cognitive Sciences* **17,** 111–120. doi:10.1016/j.tics.2013.01.001 (2013).

19. Lesburguères, E., Gobbo, O. L., Alaux-Cantin, S., Hambucken, A., Trifilieff, P. & Bontempi, B. Early Tagging of Cortical Networks Is Required for the Formation of Enduring Associative Memory. *Science* **331,** 924–928. doi:10.1126/science.1196164 (2011).

20. Richards, B. A. & Frankland, P. W. The Persistence and Transience of Memory. *Neuron* **94,** 1071–1084. doi:10.1016/j.neuron.2017.04.037 (2017).

21. Cooke, S. F., Attwell, P. J. E. & Yeo, C. H. Temporal Properties of Cerebellar-Dependent Memory Consolidation. *Journal of Neuroscience* **24,** 2934–2941. doi:10.1523/JNEUROSCI.5505-03.2004 (2004).

22. Lisberger, S. G. The Rules of Cerebellar Learning: Around the Ito Hypothesis. *Neuroscience* **462,** 164–190. doi:10.1016/j.neuroscience.2020.08.026 (2021).

23. Raymond, J. L. & Medina, J. F. Computational Principles of Supervised Learning in the Cerebellum. *Annual Review of Neuroscience* **41,** 233–253. doi:10.1146/annurev-neuro-080317-061948 (2018).

24. Andalman, A. S. & Fee, M. S. A Basal Ganglia-Forebrain Circuit in the Songbird Biases Motor Output to Avoid Vocal Errors. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 12518–12523. doi:10.1073/pnas.0903214106 (2009).

25. Warren, T. L., Tumer, E. C., Charlesworth, J. D. & Brainard, M. S. Mechanisms and Time Course of Vocal Learning and Consolidation in the Adult Songbird. *Journal of Neurophysiology* **106,** 1806–1821. doi:10.1152/jn.00311.2011 (2011).

26. Yin, H. H., Mulcare, S. P., Hilário, M. R. F., Clouse, E., Holloway, T., Davis, M. I., Hansson, A. C., Lovinger, D. M. & Costa, R. M. Dynamic Reorganization of Striatal Circuits during the Acquisition and Consolidation of a Skill. *Nature Neuroscience* **12,** 333–341. doi:10.1038/nn.2261 (2009).

27. Makino, H., Hwang, E. J., Hedrick, N. G. & Komiyama, T. Circuit Mechanisms of Sensorimotor Learning. *Neuron* **92,** 705–721. doi:10.1016/j.neuron.2016.10.029 (2016).

28. Teşileanu, T., Ölveczky, B. & Balasubramanian, V. Rules and Mechanisms for Efficient Two-Stage Learning in Neural Circuits. *eLife* **6,** e20944. doi:10.7554/eLife.20944 (2017).

29. Murray, J. M. & Escola, G. S. Remembrance of Things Practiced with Fast and Slow Learning in Cortical and Subcortical Pathways. *Nature Communications* **11,** 6441. doi:10.1038/s41467-020-19788-5 (2020).

30. Medina, J. F., Christopher Repa, J., Mauk, M. D. & LeDoux, J. E. Parallels between Cerebellum- and Amygdala-Dependent Conditioning. *Nature Reviews Neuroscience* **3,** 122–131. doi:10.1038/nrn728 (2002).

31. Broussard, D. M. & Kassardjian, C. D. Learning in a Simple Motor System. *Learning & Memory* **11,** 127–136. doi:10.1101/lm.65804 (2004).

32. Anzai, M., Kitazawa, H. & Nagao, S. Effects of Reversible Pharmacological Shutdown of Cerebellar Flocculus on the Memory of Long-Term Horizontal Vestibulo-Ocular Reflex Adaptation in Monkeys. *Neuroscience Research* **68,** 191–198. doi:10.1016/j.neures.2010.07.2038 (2010).

33. Jang, D. C., Shim, H. G. & Kim, S. J. Intrinsic Plasticity of Cerebellar Purkinje Cells Contributes to Motor Memory Consolidation. *The Journal of Neuroscience* **40,** 4145–4157. doi:10.1523/JNEUROSCI.1651-19.2020 (2020).

34. Kassardjian, C. D., Tan, Y.-F., Chung, J.-Y. J., Heskin, R., Peterson, M. J. & Broussard, D. M. The Site of a Motor Memory Shifts with Consolidation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **25,** 7979–7985. doi:10.1523/JNEUROSCI.2215-05.2005 (2005).

35. Nagao, S. & Kitazawa, H. Effects of Reversible Shutdown of the Monkey Flocculus on the Retention of Adaptation of the Horizontal Vestibulo-Ocular Reflex. *Neuroscience* **118,** 563–570. doi:10.1016/S0306-4522(02)00991-0 (2003).

36. Shutoh, F., Ohki, M., Kitazawa, H., Itohara, S. & Nagao, S. Memory Trace of Motor Learning Shifts Transsynaptically from Cerebellar Cortex to Nuclei for Consolidation. *Neuroscience* **139,** 767–777. doi:10.1016/j.neuroscience.2005.12.035 (2006).

37. Albus, J. S. A Theory of Cerebellar Function. *Mathematical Biosciences* **10,** 25–61. doi:10.1016/0025-5564(71)90051-4 (1971).

38. Clopath, C., Badura, A., De Zeeuw, C. I. & Brunel, N. A Cerebellar Learning Model of Vestibulo-Ocular Reflex Adaptation in Wild-Type and Mutant Mice. *Journal of Neuroscience* **34,** 7203–7215. doi:10.1523/JNEUROSCI.2791-13.2014 (2014).

39. Herzfeld, D. J., Hall, N. J., Tringides, M. & Lisberger, S. G. Principles of Operation of a Cerebellar Learning Circuit. *eLife* **9,** e55217. doi:10.7554/eLife.55217 (2020).

40. Ito, M. Cerebellar Control of the Vestibulo-Ocular Reflex–around the Flocculus Hypothesis. *Annual review of neuroscience* **5,** 275–297 (1982).

41. Marr, D. A Theory of Cerebellar Cortex. *The Journal of Physiology* **202,** 437–470 (1969).

42. Porrill, J. & Dean, P. Cerebellar Motor Learning: When Is Cortical Plasticity Not Enough? *PLoS Computational Biology* **3,** e197. doi:10.1371/journal.pcbi.0030197 (2007).

43. Yamazaki, T., Nagao, S., Lennon, W. & Tanaka, S. Modeling Memory Consolidation during Posttraining Periods in Cerebellovestibular Learning. *Proceedings of the National Academy of Sciences* **112,** 3541–3546. doi:10.1073/pnas.1413798112 (2015).

44. Boyden, E. S., Katoh, A. & Raymond, J. L. Cerebellum-Dependent Learning: The Role of Multiple Plasticity Mechanisms. *Annual Review of Neuroscience* **27,** 581–609. doi:10.1146/annurev.neuro.27.070203.144238 (2004).

45. Inoshita, T. & Hirano, T. Occurrence of Long-Term Depression in the Cerebellar Flocculus during Adaptation of Optokinetic Response. *eLife* **7,** e36209. doi:10.7554/eLife.36209 (2018).

46. Kimpo, R. R., Rinaldi, J. M., Kim, C. K., Payne, H. L. & Raymond, J. L. Gating of Neural Error Signals during Motor Learning. *eLife* **3,** e02076. doi:10.7554/eLife.02076 (2014).

47. McElligott, J. G., Beeton, P. & Polk, J. Effect of Cerebellar Inactivation by Lidocaine Microdialysis on the Vestibuloocular Reflex in Goldfish. *Journal of Neurophysiology* **79,** 1286–1294. doi:10.1152/jn.1998.79.3.1286 (1998).

48. Okamoto, T., Shirao, T., Shutoh, F., Suzuki, T. & Nagao, S. Post-Training Cerebellar Cortical Activity Plays an Important Role for Consolidation of Memory of Cerebellum-Dependent Motor Learning. *Neuroscience Letters* **504,** 53–56. doi:10.1016/j.neulet.2011.08.056 (2011).

49. Medina, J. F. & Mauk, M. D. Simulations of Cerebellar Motor Learning: Computational Analysis of Plasticity at the Mossy Fiber to Deep Nucleus Synapse. *The Journal of Neuroscience* **19,** 7140–7151. doi:10.1523/JNEUROSCI.19-16-07140.1999 (1999).

50. Menzies, J. R. W., Porrill, J., Dutia, M. & Dean, P. Synaptic Plasticity in Medial Vestibular Nucleus Neurons: Comparison with Computational Requirements of VOR Adaptation. *PLoS ONE* **5** (ed Tell, F.) e13182. doi:10.1371/journal.pone.0013182 (2010).

51. Burak, Y. & Fiete, I. R. Fundamental Limits on Persistent Activity in Networks of Noisy Neurons. *Proceedings of the National Academy of Sciences* **109,** 17645–17650. doi:10.1073/pnas.1117386109 (2012).

52. Ganguli, S., Huh, D. & Sompolinsky, H. Memory Traces in Dynamical Systems. *Proceedings of the National Academy of Sciences* **105,** 18970–18975. doi:10.1073/pnas.0804451105 (2008).

53. Lim, S. & Goldman, M. S. Noise Tolerance of Attractor and Feedforward Memory Models. *Neural Computation* **24,** 332–390. doi:10.1162/NECO_a_00234 (2012).

54. Dean, P., Porrill, J. & Stone, J. V. Decorrelation Control by the Cerebellum Achieves Oculomotor Plant Compensation in Simulated Vestibulo-Ocular Reflex. *Proceedings of the Royal Society of London B: Biological Sciences* **269,** 1895–1904. doi:10.1098/rspb.2002.2103 (2002).

55. Miller, K. D. & MacKay, D. J. C. The Role of Constraints in Hebbian Learning. *Neural Computation* **6,** 100–126. doi:10.1162/neco.1994.6.1.100 (1994).

56. Loewenstein, Y. Robustness of Learning That Is Based on Covariance-Driven Synaptic Plasticity. *PLoS Computational Biology* **4,** e1000007. doi:10.1371/journal.pcbi.1000007 (2008).

57. Oja, E. Simplified Neuron Model as a Principal Component Analyzer. *Journal of Mathematical Biology* **15,** 267–273. doi:10.1007/BF00275687 (1982).

58. Bienenstock, E. L., Cooper, L. N. & Munro, P. W. Theory for the Development of Neuron Selectivity: Orientation Specificity and Binocular Interaction in Visual Cortex. *Journal of Neuroscience* **2,** 32–48. doi:10.1523/JNEUROSCI.02-01-00032.1982 (1982).

59. Chistiakova, M., Bannon, N. M., Chen, J.-Y., Bazhenov, M. & Volgushev, M. Homeostatic Role of Heterosynaptic Plasticity: Models and Experiments. *Frontiers in Computational Neuroscience* **9.** doi:10.3389/fncom.2015.00089 (2015).

60. Turrigiano, G. G. The Self-Tuning Neuron: Synaptic Scaling of Excitatory Synapses. *Cell* **135,** 422–435. doi:10.1016/j.cell.2008.10.008 (2008).

61. Yger, P. & Gilson, M. Models of Metaplasticity: A Review of Concepts. *Frontiers in Computational Neuroscience* **9,** 138. doi:10.3389/fncom.2015.00138 (2015).

62. Miles, F. A. & Lisberger, S. G. Plasticity in the Vestibulo-Ocular Reflex: A New Hypothesis. *Annual review of neuroscience* **4,** 273–299 (1981).

63. Lisberger, S. G. & Sejnowski, T. J. Motor Learning in a Recurrent Network Model Based on the Vestibulo-Ocular Reflex. *Nature* **360,** 159–161 (1992).

64. Lisberger, S. G. Neural Basis for Motor Learning in the Vestibuloocular Reflex of Primates. III. Computational and Behavioral Analysis of the Sites of Learning. *Journal of Neurophysiology* **72,** 974–998 (1994).

65. Person, A. L. Corollary Discharge Signals in the Cerebellum. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **4,** 813–819. doi:10.1016/j.bpsc.2019.04.010 (2019).

66. Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C. & Gerstner, W. Inhibitory Plasticity Balances Excitation and Inhibition in Sensory Pathways and Memory Networks. *Science* **334,** 1569–1573. doi:10.1126/science.1211095 (2011).

67. Miles, F. A., Braitman, D. J. & Dow, B. M. Long-Term Adaptive Changes in Primate Vestibuloocular Reflex. IV. Electrophysiological Observations in Flocculus of Adapted Monkeys. *Journal of Neurophysiology* **43,** 1477–1493 (1980).

68. Payne, H. L., Raymond, J. L. & Goldman, M. S. Interactions between Circuit Architecture and Plasticity in a Closed-Loop Cerebellar System. *eLife* **13,** e84770. doi:10.7554/eLife.84770 (2024).

69. Churchland, A. K. & Ditterich, J. New Advances in Understanding Decisions among Multiple Alternatives. *Current Opinion in Neurobiology* **22,** 920–926. doi:10.1016/j.conb.2012.04.009 (2012).

70. Gold, J. I. & Shadlen, M. N. The Neural Basis of Decision Making. *Annual Review of Neuroscience* **30,** 535–574. doi:10.1146/annurev.neuro.29.051605.113038 (2007).

71. Usher, M. & McClelland, J. L. The Time Course of Perceptual Choice: The Leaky, Competing Accumulator Model. *Psychological Review* **108,** 550–592. doi:10.1037/0033-295X.108.3.550 (2001).

72. Wang, X.-J. Decision Making in Recurrent Neuronal Circuits. *Neuron* **60,** 215–234. doi:10.1016/j.neuron.2008.09.034 (2008).

73. Cannon, S. C., Robinson, D. A. & Shamma, S. A Proposed Neural Network for the Integrator of the Oculomotor System. *Biological Cybernetics* **49,** 127–136. doi:10.1007/BF00320393 (1983).

74. Goldman, M. S., Levine, J. H., Major, G., Tank, D. W. & Seung, H. S. Robust Persistent Neural Activity in a Model Integrator with Multiple Hysteretic Dendrites per Neuron. *Cerebral Cortex* **13,** 1185–1195. doi:10.1093/cercor/bhg095 (2003).

75. Koulakov, A. A., Raghavachari, S., Kepecs, A. & Lisman, J. E. Model for a Robust Neural Integrator. *Nature Neuroscience* **5,** 775–782. doi:10.1038/nn893 (2002).

76. O'Connor, D. H., Wittenberg, G. M. & Wang, S. S.-H. Graded Bidirectional Synaptic Plasticity Is Composed of Switch-like Unitary Events. *Proceedings of the National Academy of Sciences* **102,** 9679–9684. doi:10.1073/pnas.0502332102 (2005).

77. Nikitchenko, M. & Koulakov, A. Neural Integrator: A Sandpile Model. *Neural Computation* **20,** 2379–2417. doi:10.1162/neco.2008.12-06-416 (2008).

78. Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U. & Nieuwenhuis, S. The Neural Basis of the Speed–Accuracy Tradeoff. *Trends in Neurosciences* **33,** 10–16. doi:10.1016/j.tins.2009.09.002 (2010).

79. Alvarez, P. & Squire, L. R. Memory Consolidation and the Medial Temporal Lobe: A Simple Network Model. *Proceedings of the National Academy of Sciences* **91,** 7041–7045. doi:10.1073/pnas.91.15.7041 (1994).

80. Remme, M. W. H., Bergmann, U., Alevi, D., Schreiber, S., Sprekeler, H. & Kempter, R. Hebbian Plasticity in Parallel Synaptic Pathways: A Circuit Mechanism for Systems Memory Consolidation. *PLOS Computational Biology* **17** (ed Bush, D.) e1009681. doi:10.1371/journal.pcbi.1009681 (2021).

81. Tomé, D. F., Sadeh, S. & Clopath, C. Coordinated Hippocampal-Thalamic-Cortical Communication Crucial for Engram Dynamics underneath Systems Consolidation. *Nature Communications* **13,** 840. doi:10.1038/s41467-022-28339-z (2022).

82. Wittenberg, G. M., Sullivan, M. R. & Tsien, J. Z. Synaptic Reentry Reinforcement Based Network Model for Long-Term Memory Consolidation. *Hippocampus* **12,** 637–647. doi:10.1002/hipo.10102 (2002).

83. Dong, D. W. & Hopfield, J. J. Dynamic Properties of Neural Networks with Adapting Synapses. *Network: Computation in Neural Systems* **3,** 267–283. doi:10.1088/0954-898X_3_3_002 (1992).

84. Dayan, P. & Abbott, L. F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, MA, 2005).

85. Kuśmierz, Ł., Isomura, T. & Toyoizumi, T. Learning with Three Factors: Modulating Hebbian Plasticity with Errors. *Current Opinion in Neurobiology* **46,** 170–177. doi:10.1016/j.conb.2017.08.020 (2017).

86. Urbanczik, R. & Senn, W. Learning by the Dendritic Prediction of Somatic Spiking. *Neuron* **81,** 521–528. doi:10.1016/j.neuron.2013.11.030 (2014).

87. Johnson, R. & Zhang, T. *Accelerating Stochastic Gradient Descent Using Predictive Variance Reduction* in *Advances in Neural Information Processing Systems* **26** (2013).

88. Sekeres, M. J., Winocur, G. & Moscovitch, M. The Hippocampus and Related Neocortical Structures in Memory Transformation. *Neuroscience Letters* **680,** 39–53. doi:10.1016/j.neulet.2018.05.006 (2018).

89. Tse, D., Langston, R. F., Kakeyama, M., Bethus, I., Spooner, P. A., Wood, E. R., Witter, M. P. & Morris, R. G. M. Schemas and Memory Consolidation. *Science* **316,** 76–82. doi:10.1126/science.1135935 (2007).

90. Lindsey, J. & Litwin-Kumar, A. *Theory of Systems Memory Consolidation via Recall-Gated Plasticity* preprint (eLife, 2023). doi:10.7554/eLife.90793.1.

91. Sun, W., Advani, M., Spruston, N., Saxe, A. & Fitzgerald, J. E. Organizing Memories for Generalization in Complementary Learning Systems. *Nature Neuroscience.* doi:10.1038/s41593-023-01382-9 (2023).

92. Jacquerie, K., Minne, C., Ponnet, J., Benghalem, N., Sacré, P. & Drion, G. *Switches to Slow Rhythmic Neuronal Activity Lead to a Plasticity-Induced Reset in Synaptic Weights* preprint (bioRxiv, 2023). doi:10.1101/2022.07.15.500198.

93. Leimer, P., Herzog, M. & Senn, W. *Synaptic Weight Decay with Selective Consolidation Enables Fast Learning without Catastrophic Forgetting* preprint (bioRxiv, 2019). doi:10.1101/613265.

94. Li, H. L. & Van Rossum, M. C. Energy Efficient Synaptic Plasticity. *eLife* **9,** e50804. doi:10.7554/eLife.50804 (2020).

95. Körding, K. P., Tenenbaum, J. B. & Shadmehr, R. The Dynamics of Memory as a Consequence of Optimal Adaptation to a Changing Body. *Nature Neuroscience* **10,** 779–786. doi:10.1038/nn1901 (2007).

96. Carr, M. F., Jadhav, S. P. & Frank, L. M. Hippocampal Replay in the Awake State: A Potential Substrate for Memory Consolidation and Retrieval. *Nature Neuroscience* **14,** 147–153. doi:10.1038/nn.2732 (2011).

97. Murre, J. M. TraceLink: A Model of Amnesia and Consolidation of Memory. *Hippocampus* **6,** 675–684. doi:10.1002/(SICI)1098-1063(1996)6:6<675::AID-HIPO10>3.0.CO;2-Y (1996).

98. Pugh, J. R. & Raman, I. M. Potentiation of Mossy Fiber EPSCs in the Cerebellar Nuclei by NMDA Receptor Activation Followed by Postinhibitory Rebound Current. *Neuron* **51,** 113–123. doi:10.1016/j.neuron.2006.05.021 (2006).

99. Hirano, T. Regulation and Interaction of Multiple Types of Synaptic Plasticity in a Purkinje Neuron and Their Contribution to Motor Learning. *The Cerebellum* **17,** 756–765. doi:10.1007/s12311-018-0963-0 (2018).

100. Mapelli, L., Pagani, M., Garrido, J. A. & D'Angelo, E. Integrated Plasticity at Inhibitory and Excitatory Synapses in the Cerebellar Circuit. *Frontiers in Cellular Neuroscience* **9.** doi:10.3389/fncel.2015.00169 (2015).

101. Najac, M. & Raman, I. M. Integration of Purkinje Cell Inhibition by Cerebellar Nucleo-Olivary Neurons. *Journal of Neuroscience* **35,** 544–549. doi:10.1523/JNEUROSCI.3583-14.2015 (2015).

102. Uusisaari, M. & Knöpfel, T. Functional Classification of Neurons in the Mouse Lateral Cerebellar Nuclei. *The Cerebellum* **10,** 637–646. doi:10.1007/s12311-010-0240-3 (2011).

103. Fanning, A. S., Shakhawat, A. M. & Raymond, J. L. Population Calcium Responses of Purkinje Cells in the Oculomotor Cerebellum Driven by Nonvisual Input. *Journal of Neurophysiology* **126,** 1391–1402. doi:10.1152/jn.00715.2020 (2021).

104. Winkelman, B. & Frens, M. Motor Coding in Floccular Climbing Fibers. *Journal of Neurophysiology* **95,** 2342–2351. doi:10.1152/jn.01191.2005 (2006).

Materials and methods

Code for simulations, analyses and producing plots for figures can be found on GitHub at: https://github.com/goldman-lab/consolidation-integration.

Feedforward sensorimotor circuit model

Here we describe the feedforward model of Figures 2–4. The circuit model has parallel pathways that transform a graded, time-varying input into a response. In particular, we modeled the transformation of a graded, time-varying sensory (vestibular) input into an eye movement response and its modification by cerebellum-dependent learning. We modeled each node in the circuit with a single variable representing the average firing rate of a population of cells: mossy fibers, $\mathrm{MF}(t)$; parallel fibers (granule cell axons), $\mathrm{PF}(t)$; Purkinje cells, $\mathrm{PC}(t)$; medial vestibular nucleus neurons, $\mathrm{MVN}(t)$; and climbing fibers, $\mathrm{CF}(t)$ (Fig. 2*A*). Specific values of model parameters used in simulations were taken from the literature, where available, and are summarized in Table 1. For analysis, unless otherwise stated, all parameters are assumed to be nonnegative.

The firing rate of mossy fibers was modeled as a combination of a spontaneous baseline and a time-varying component encoding a sensory (vestibular) input driven by head motion $\dot{H}(t)$,

$$\mathrm{MF}(t) = \mathrm{MF}_0 + \delta\mathrm{MF}(t) = \mathrm{MF}_0 + k_{\mathrm{MF}}\dot{H}(t), \tag{1}$$

where we considered the encoding of the sensory input to be linear with sensitivity to head velocity k_{MF} (Lasker *et al.* 2008). Rather than explicitly model the mossy fiber to granule cell transformation, we considered the firing rate of parallel fibers to similarly depend on vestibular input linearly,

$$\mathrm{PF}(t) = \mathrm{PF}_0 + \delta\mathrm{PF}(t) = \mathrm{PF}_0 + k_{\mathrm{PF}}\dot{H}(t), \tag{2}$$

where k_{PF} was the firing rate sensitivity to head velocity.

Purkinje cells, in the early-learning area, were modeled as linearly combining direct excitatory input from parallel fibers and indirect inhibitory input via a parallel fiber-molecular layer interneuron pathway (Yamazaki *et al.* 2015),

$$\mathrm{PC}(t) = \mathrm{PC}_0 + (w_H^+(t) - w_H^-)\mathrm{PF}(t) = \mathrm{PC}_0 + w_H(t)\mathrm{PF}(t), \tag{3}$$

where w_H^+ and w_H^- are the total excitatory and inhibitory weights onto the Purkinje cell, so the early-learning weight $w_H = w_H^+ - w_H^-$ is the net weight of the parallel fiber pathway onto Purkinje cells. For simplicity, the inhibitory weight was not plastic in the model; the dynamics of learning are driven by the net weight. Equivalently, $w_H(t)$ can be interpreted simply as modeling the dynamics of the net (excitatory minus inhibitory) weight onto Purkinje cells. The effect of climbing fibers, which fire at very low rates (~ 1 sp/s), on Purkinje cell output was not modeled here.

The medial vestibular nucleus, the late-learning area, was modeled as responding linearly to its excitatory mossy fiber input and inhibitory Purkinje cell input (Bagnall *et al.* 2008; Beraneck & Cullen 2007),

$$\begin{aligned}\mathrm{MVN}(t) &= \mathrm{MVN}_0 + v(t)\mathrm{MF}(t) - w_{\mathrm{PC}}\mathrm{PC}(t) \\ &= \mathrm{MVN}_0 - w_{\mathrm{PC}}\mathrm{PC}_0 + v(t)\mathrm{MF}(t) - w_H(t)w_{\mathrm{PC}}\mathrm{PF}(t),\end{aligned} \tag{4}$$

where v is the late-learning weight of the vestibular mossy fiber input to the medial vestibular nucleus, w_{PC} is the weight of the Purkinje cell input to the MVN, and MVN_0 is an offset chosen so that the spontaneous MVN firing rate matched experimental values (57 sp/s) (Beraneck & Cullen 2007). The initial value of v before training was chosen so that the gain in the dark was 0.4 (Kimpo *et al.* 2014; Boyden & Raymond 2003).

Eye movement output, $\dot{E}(t)$, was modeled as being proportional to the deviation of medial vestibular nucleus neuron firing from its baseline (Yamazaki *et al.* 2015; Clopath *et al.* 2014),

$$\dot{E}(t) = -k_E(\mathrm{MVN}(t) - \langle \mathrm{MVN}(t)\rangle_{\tau_f}). \tag{5}$$

The baseline, $\langle \mathrm{MVN}(t)\rangle_{\tau_f}$, was calculated as an exponential (low-pass filtered) average over the recent past of MVN activity. For any time-varying quantity $x(t)$, its exponential average $\langle x(t)\rangle_\tau$ over timescale τ is defined by

$$\tau\frac{\mathrm{d}\langle x\rangle_\tau}{\mathrm{d}t} = -\langle x\rangle_\tau + x(t). \tag{6}$$

The timescale of the average $\tau_f = 0.017$ h (1 min) is long relative to variations in the sensory input, but is much shorter than the timescales of plasticity (defined below), so that $\langle w_H(t)\rangle_{\tau_f} \approx w_H(t)$ and $\langle v(t)\rangle_{\tau_f} \approx v(t)$ (Yamazaki *et al.* 2015). Throughout the paper, we assume that over the timescale τ_f, the sensory input is zero on average,

$$\langle \dot{H}(t)\rangle_{\tau_f} \approx 0. \tag{7}$$

In the specific case of the bidirectional vestibular input to the oculomotor circuit, this corresponds to the time average of the rightward and leftward head movements being approximately equal.

The input-to-output gain of the circuit, g, is defined as the ratio of eye velocity output to head velocity input,

$$g(t) = -\frac{\dot{E}(t)}{\dot{H}(t)} \approx k_E(k_{\mathrm{MF}}v(t) - k_{\mathrm{PF}}w_{\mathrm{PC}}w_H(t)), \tag{8}$$

where the negative sign is present because eye movements are oppositely directed from head movements to keep the image of the world stable.

Learning rules

Plasticity at the early-learning site was driven by feedback about behavioral errors, carried by the climbing fibers. Climbing fiber firing was represented by the sum of a spontaneous baseline rate and a time-varying component representing oculomotor errors,

$$\mathrm{CF}(t) = \mathrm{CF}_0 + \delta\mathrm{CF}(t). \tag{9}$$

The time-varying component $\delta\mathrm{CF}(t)$ encodes retinal slip, or a failure of eye movements to stabilize images on the retina, which occurs when the input-output gain of the circuit is incorrect. During training, the target output $\dot{E}^{\mathrm{target}}$ of the circuit is a (negatively) scaled version of the vestibular input $\dot{H}(t)$ with gain g^{target},

$$\dot{E}^{\mathrm{target}}(t) = -g^{\mathrm{target}}\dot{H}(t), \tag{10}$$

so, using Eq. (8), the retinal slip error is

$$\dot{R}(t) = \dot{E}^{\mathrm{target}}(t) - \dot{E}(t) = -(g^{\mathrm{target}} - g(t))\dot{H}(t). \tag{11}$$

We model the time-varying component of the climbing fiber firing as

$$\delta\mathrm{CF}(t) = k_{\mathrm{CF}} \tanh(-\beta_{\mathrm{light}}\dot{R}), \tag{12}$$

which is a saturating function of retinal slip. We chose the magnitude of the saturation to be $k_{\mathrm{CF}} = \mathrm{CF}_0 = 1$ Hz, so that when retinal slip errors are large, climbing fiber firing saturates at a maximum of 2 sp/s and a minimum of 0 sp/s (Guo *et al.* 2014; Winkelman & Frens 2006). When retinal slip errors are relatively small (relative to $1/\beta_{\mathrm{light}}$), $\delta\mathrm{CF}(t)$ is approximately linearly related to negative retinal slip. Thus, if the eye movements are too small so that the gain of the response needs to be increased, the covariance between the climbing fiber and parallel fiber firing rates is positive.

Experimentally, coactivation of parallel fibers and climbing fibers causes associative long-term depression at the early-learning site, whereas activation of parallel fibers alone causes long-term potentiation (Raymond & Lisberger 1998; Suvrathan *et al.* 2016; Lev-Ram *et al.* 2002). Therefore, plasticity at the early-learning site was modeled as

$$\tau_w \frac{\mathrm{d}w_H}{\mathrm{d}t} = -(w_H(t) + w_H^-) + k_{\mathrm{LTP}}\langle \mathrm{PF}(t)\rangle_{\tau_{f,w}} - k_{\mathrm{LTD}}\langle \mathrm{PF}(t)\cdot \mathrm{CF}(t)\rangle_{\tau_{f,w}}, \tag{13}$$

where the first term on the right is a passive decay, or forgetting, in the value of w_H^+ (Yamazaki *et al.* 2015; Clopath *et al.* 2014). We chose k_{LTP} and k_{LTD} so that the steady state value of w_H^+ when not training was equal to w_H^- (see *Analysis of synaptic learning dynamics*). The time constant of plasticity, τ_w, was chosen to be relatively fast (0.15 h = 9 min) during training (Boyden & Raymond 2003) and slower (5 h) during the post-training period (Jang *et al.* 2020; Cooke *et al.* 2004). The

timescale of the averages over inputs, $\tau_{f,w}$, was chosen to be the same as τ_f.

We considered two candidate learning rules for the late-learning site: a heterosynaptic and a Hebbian covariance-like rule. Heterosynaptic plasticity was modeled as

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -k_{v,\mathrm{hetero}}\langle \mathrm{MF}(t)(\mathrm{PC}(t) - \mathrm{PC}_0)\rangle_{\tau_{f,v}}, \tag{14}$$

where an increase in v is driven by the anticorrelation between mossy fiber firing and fluctuations of Purkinje cell activity around baseline (Herzfeld *et al.* 2020; Clopath *et al.* 2014; Porrill & Dean 2007; Medina & Mauk 1999; Menzies *et al.* 2010; Pugh & Raman 2006). Hebbian plasticity was modeled as

$$\frac{\mathrm{d}v}{\mathrm{d}t} = k_{v,\mathrm{Hebb}}\langle \mathrm{MF}(t)(\mathrm{MVN}(t) - \theta(t))\rangle_{\tau_{f,v}}, \tag{15}$$

$$\tau_s \frac{\mathrm{d}\theta}{\mathrm{d}t} = -\theta + \mathrm{MVN}(t), \tag{16}$$

where the synaptic dynamics consist of the "readout" weight v and an internal sliding threshold variable θ that calculates a leaky average of postsynaptic activity (Yamazaki *et al.* 2015). The threshold is similar to those used in previous studies to counteract the runaway plasticity that would otherwise result from the inherent positive feedback driven by this kind of correlational rule (Zenke *et al.* 2017; Miller & MacKay 1994; Bienenstock *et al.* 1982; see SI Text, S2). For stability, the timescale of the sliding threshold was chosen to be fast relative to the rate of plasticity (see SI Text, S1 for details; Zenke *et al.* (2013)).

Simulation of oculomotor learning

We simulated oculomotor experiments with 0.5 h of training to increase the gain of the eye movement response, followed by a 23.5 h post-training period. During training, vestibular input was sinusoidal,

$$\dot{H}(t) = v_{\mathrm{peak}} \sin(2\pi f t), \tag{17}$$

with peak velocity $v_{\mathrm{peak}} = 15$ deg/s and rotational frequency $f = 1$ Hz, which is representative of experimental protocols (Jang *et al.* 2020; Kimpo *et al.* 2014; Nguyen-Vu *et al.* 2013; Boyden & Raymond 2003). In all simulations, we chose parameters such that the initial gain of the response was 0.4. In the simulations shown in Figure 2, we used a target gain value of $g^{\mathrm{target}} = 2$, and chose w_H^-, k_{LTP} and k_{LTD} so that during training, w_H^+ decreased by 51% (Jang *et al.* 2020) and the input-to-output gain increased by 30% from baseline (Kimpo *et al.* 2014; Boyden & Raymond 2003). To understand the stability of the dynamics during training, we also simulated a single 48 h training period to increase the gain to $g^{\mathrm{target}} = 1$ (Fig. S1; see *Analysis of synaptic learning dynamics*). All other model parameters are shown in Table 1.

During the post-training period, we set $\dot{R} \equiv 0$. This models the animal being placed in the dark

after training to eliminate feedback about oculomotor errors (Anzai *et al.* 2010; Shutoh *et al.* 2006; Kassardjian *et al.* 2005; Nagao & Kitazawa 2003; Boyden & Raymond 2003). We considered both the case in which vestibular input was not present post-training (as in Yamazaki *et al.* 2015), as well as the case in which it was present (as in Clopath *et al.* 2014). In the latter case, we used the same vestibular input as during training. For the heterosynaptic rule, the presence of post-training input caused a slightly larger increase in v than when input was absent (not shown), unlike for the Hebbian rule (Fig. 5). The learning rates $k_{v,\mathrm{hetero}}$ and $k_{v,\mathrm{Hebb}}$ were picked so that approximately 75% of the increase in gain during the training period was consolidated (Boyden & Raymond 2003).

For the heterosynaptic rule, we studied the robustness of consolidated memories to perturbations in the early-learning site by adding random values to the early-learning weight w_H at regular intervals (Fig. 3*A*,*B*). We initialized the weights to the same baseline as the training simulations and, every 10 minutes of simulation time, added a value chosen uniformly at random between -1.7 and 1.7 to w_H. For the simulation shown in Fig. 3*A*, we used the same parameters as for the simulation in Fig. 2 (Table 1), but with $\tau_{f,v} = \tau_f$. For the simulation shown in Fig. 3*B*, we used a smaller learning rate, $k_{v,\mathrm{hetero}} = 7.73 \times 10^{-6}\ (\mathrm{s/sp})^2/\mathrm{h}$. We performed 250 runs of the perturbation simulation with the normal and smaller values of $k_{v,\mathrm{hetero}}$, with each simulation run for 10 hours of simulation time. The histograms in Fig. 3*C* were constructed from the endpoints of these 250 simulations, and we plotted the time course of the variances of the corresponding distributions (at the time point immediately before a perturbation) in Fig. 3*D*. In Fig. 3*E* and *F*, we simulated training using the same parameters as for the perturbation simulations in panels *A* and *B* respectively.

All simulations were performed in Python by integrating the differential equations for the learning rules with the Radau solver (implemented in `solve_ivp` in the `scipy.integrate` package). Time steps in the simulation had a base unit of hours.

Analysis of synaptic learning dynamics

To understand the conditions for successful consolidation in the model, we examined the dynamics of learning analytically. To simplify our analysis, we consider timescales that are relatively long compared to variation in the sensory input, so that the input is zero on average, Eq. (7), and the variability of the sensory input, as measured by the time average of the squared input, is constant,

$$\langle \dot{H}(t)^2 \rangle_{\tau_f} \approx \sigma_H^2. \tag{18}$$

Using these assumptions, the dynamics of the early-learning site can be described by simplifying Eq. (13) to

$$\tau_w \frac{\mathrm{d}w_H}{\mathrm{d}t} \approx -(w_H(t) + w_H^-) + w_\infty^+ - c(t), \tag{19}$$

where $w_\infty^+ = \mathrm{PF}_0(k_{\mathrm{LTP}} - k_{\mathrm{LTD}}\mathrm{CF}_0)$ depends only on the baseline firing rates of inputs to the Purkinje cell, and $c(t) = k_{\mathrm{LTD}}\langle \delta\mathrm{PF}(t)\delta\mathrm{CF}(t)\rangle_{\tau_f}$ is proportional to the covariance of the parallel

fiber-mediated sensory input and climbing fiber feedback about errors. When the circuit is producing behavioral output without error (as we assume is the case before training), or during the post-training period, when there is no feedback about errors, $c(t) \equiv 0$. As a result, w_H tends toward a baseline value of $w_\infty^+ - w_H^-$. If $c(t) > 0$ during training, as was the case in our simulations, w_H decreases from this baseline, increasing the input-output gain of the circuit at the end of the training period. On the other hand, eye movements that are too large would result in $c(t) < 0$, causing an increase in w_H and a decrease in the gain.

At the late-learning site, the heterosynaptic rule, Eq. (14), can be simplified to

$$\frac{\mathrm{d}v}{\mathrm{d}t} \approx -k_{v,\mathrm{hetero}}(\mathrm{MF}_0 \cdot \mathrm{PF}_0 + k_{\mathrm{MF}} k_{\mathrm{PF}} \sigma_H^2) w_H(t), \tag{20}$$

where we assume that the timescale of plasticity of w_H is significantly longer than τ_f (the timescale of correlations in neural activity to which the learning rule is sensitive) and therefore the value of w_H is effectively uncorrelated with the sensory input. From the form of Eq. (20), the late-learning weight v is proportional to the temporal integral of the early-learning weight w_H.

We can then understand the dynamics of learning as follows. Initially, when errors are large and correlated with the sensory input, because of the saturation in the CF response, we have

$$c(t) \approx k_{\mathrm{LTD}} k_{\mathrm{PF}} k_{\mathrm{CF}} C, \tag{21}$$

where

$$C = \langle |\dot{H}| \rangle_{\tau_{f,w}}. \tag{22}$$

For sinusoidal sensory input with period of oscillation much shorter than $\tau_{f,w}$, $C \approx 2/\pi \cdot v_{\mathrm{peak}}$, which is constant for a fixed choice of peak input amplitude. During training, when $c(t) \neq 0$, w_H tends toward a value that is smaller than baseline by an amount proportional to C (Fig. S1*A,B*). This saturation can be understood as long-term depression driven by error being balanced by the decay in w_H to baseline. As v starts to integrate, this further decreases the magnitude of error. When errors are decreased enough such that $c(t)$ is approximately linear in terms of error magnitude,

$$c(t) \approx k_{\mathrm{LTD}} k_{\mathrm{PF}} k_{\mathrm{CF}} \beta_{\mathrm{light}} (g^{\mathrm{target}} - g(t)) \sigma_H^2. \tag{23}$$

In this regime, there is a stable fixed point of the weights that corresponds to the target input-to-output gain (Fig. S1*B,C*; see below and SI Text, S3).

After training ends, w_H returns to baseline. v must also go to a steady state in order for the system to be stable. From Eq. (20), this occurs as long as w_H has a baseline value of 0, which from Eq. (19) requires that

$$w_\infty^+ = w_H^-, \tag{24}$$

that is, when at steady state the excitatory and inhibitory weights of the sensory input to the

early-learning site are balanced. As a result, there is a continuum of steady states of the system in the post-training period, corresponding to a continuum of input-to-output gains of the circuit, since any weight configuration where $w_H = 0$ is a fixed point of the dynamics, which can be represented as a line in w_H-v space (Fig. 2*E*). We note that the stable fixed points of the system during training, as determined by the target gain g^{target}, also lie along the same line (Fig. S1*C*). When the steady-state excitation-inhibition balance condition, Eq. (24), is met, any initial set of weights will evolve toward some fixed point on the line $w_H = 0$. Therefore, $w_H = 0$ is furthermore a line attractor. This can be visualized by plotting a flow field in w_H-v space (Fig. 2*D,E*), where a vector describing the direction of a trajectory at any point is given by $\mathbf{x}(w_H, v) = [\mathrm{d}w_H/\mathrm{d}t, \mathrm{d}v/\mathrm{d}t]$.

The overall dynamics of learning and consolidation can thus be put succinctly: oculomotor errors drive a change Δw_H in the early-learning weight that is temporally integrated into a persistent change Δv in the late-learning weight, as the change in w_H decays away. The need to transform a negative Δw_H into a positive Δv explains why plasticity at v is driven by the *anti*correlation between mossy fiber and Purkinje cell fluctuations in Eq. (14).

Using the flow field, we can understand how the learning rates at the early- and late-learning sites control the amount of consolidation after a period of training. We assume that the circuit starts in a consolidated state before training, so that during the training period w_H changes from its baseline value of 0 by an amount Δw_H. We also assume that v does not change much during the training period. From Eqs. (19) and (20), the slope of the vectors $\mathbf{x}(w_H, v)$ describing the flow during the post-training period is

$$\frac{\mathrm{d}v}{\mathrm{d}w_H} \approx k_{v,\text{hetero}}\tau_w(\mathrm{MF}_0 \cdot \mathrm{PF}_0 + k_{\mathrm{MF}}k_{\mathrm{PF}}\sigma_H^2). \tag{25}$$

In other words, geometrically, the weight configuration moves approximately along a line with slope given by Eq. (25), so that as w_H returns to 0, the change in v during the post-training period is given by $\Delta v \approx -\mathrm{d}v/\mathrm{d}w_H \cdot \Delta w_H$. Since the slope is proportional to $k_{v,\text{hetero}}\tau_w$, a learning rate $k_{v,\text{hetero}}$ at v that is much smaller than the learning rate $1/\tau_w$ at w_H would correspond to slow consolidation, but robustness of v to changes in w_H during the post-training period (Fig. 3*B,F*).

For the circuit to completely consolidate the changes made at the early-learning site during the training period—i.e., preserve the value of the gain g_{train} achieved during training—trajectories need to follow the line in synaptic weight space corresponding to a constant gain (Eq. (8), with $g(t) \equiv g_{\text{train}}$). This is achieved if the slope of the flow field is close to that of the constant gain line,

$$k_{v,\text{hetero}}\tau_w\langle \mathrm{MF} \cdot \mathrm{PF}\rangle_{\tau_{f,v}} \approx \frac{w_{\mathrm{PC}}k_{\mathrm{PF}}}{k_{\mathrm{MF}}}, \tag{26}$$

(Fig. 3*A,B*, *top*, dashed lines). In general, we can write the fraction p of the gain change induced

during training that is consolidated post-training as

$$p = \frac{g_{\text{cons}} - g_{\text{pre}}}{g_{\text{train}} - g_{\text{pre}}} = \frac{k_{v,\text{hetero}}\tau_w \langle \text{MF} \cdot \text{PF} \rangle_{\tau_{f,v}}}{w_{\text{PC}} \cdot k_{\text{PF}}/k_{\text{MF}}}, \tag{27}$$

where g_{pre} is the gain before training, g_{train} is the gain immediately after training, and g_{cons} is the gain after consolidation. Complete consolidation (i.e., Eq. (26)) corresponds to $p = 1$.

As noted in the main text, because v is integrating, noise is also accumulated. We examined this by perturbing w_H in the absence of training signals from feedback about errors (see *Simulation of oculomotor learning*). Extending the current analysis to the case of random perturbations, we can show that, for a sequence of perturbations in w_H that are uniformly distributed over $[-\alpha, \alpha]$, the variance across simulations of the value of v after the kth perturbation will be

$$\text{Var}(v^{(k)}) = \frac{1}{3}\alpha^2 (k_{v,\text{hetero}}\tau_w \langle \text{MF} \cdot \text{PF} \rangle_{\tau_{f,v}})^2 (k + 1 + A(k)), \tag{28}$$

where $A(k)$ is a function of the number of perturbations and the time between perturbations (see SI Text, S4, for definition and derivation). We can see that the drift in the memory due to noise (as measured by the variance, Eq. (28)) and the fraction of the memory consolidated post-training, Eq. (27), are both increasing functions of the slope of the flow field, Eq. (25), leading to a tradeoff between the robustness of the memory to drift and the speed at which learning is consolidated (Fig. 3*G*).

The slope of the flow field is also related to the variability of sensory input σ_H^2. In this way, an input with larger variability would increase the rate of plasticity at v. However, this contribution to the rate of plasticity due to stimulus-driven variability is much smaller than that due to the high baseline firing rates of the mossy fibers MF_0 and parallel fibers PF_0 (not shown; see Eq. (25)) given the parameters we used, which were chosen to match experimental values.

A similar analysis for the Hebbian covariance-like rule shows that a line attractor may exist when there is no variability in the sensory input, $\sigma_H^2 = 0$ (Fig. 5*A,B*, light lines), but that the introduction of variability collapses the line attractor into a single unstable fixed point, which would not allow the circuit to learn an analog input-to-output gain (Fig. 5*A,B*, dark lines; SI Text, S1). As discussed in the main text (*Implications for plasticity rules*), simple implementations of classic homeostatic Hebbian learning rules that fix the norm of the weight(s) onto the neuron (Miller & MacKay 1994; Oja 1982) or attempt to maintain postsynaptic firing rate at a target value (Yger & Gilson 2015; Bienenstock *et al.* 1982) also failed to support a continuum of stable weight configurations in the model (see SI Text, S2). Our analysis found that a modified Hebbian covariance rule containing a decay could stably support a continuum of steady states even for $\sigma_H^2 > 0$, but this required meeting a fine-tuning condition that depends on the value of σ_H^2 (see SI Text, S1).

Mitigation of stability-plasticity dilemma with two sites of learning

Here we outline some simple analyses we performed to understand how the stability-plasticity dilemma is addressed in the presence of noisy instructive signals driving learning in the circuit. We modeled learning across multiple training sessions where during each session k, the target gain $g^{\text{target}} = \hat{g}^{(k)}$ used to calculate retinal slip was randomly drawn from a normal distribution. Target gain values were drawn independently across training sessions. In the first section below, we show that consolidation, as an integration process, causes the circuit to effectively average over the noise in the instructive signal. In the next section, we show that averaging allows the circuit to mitigate a stability-plasticity tradeoff that would occur if the circuit only had one site of plasticity.

Consolidation averages over variability in instructive signals For the analytic calculations below, we use a simplified version of the model that ignores the exact time course of the weight dynamics, and instead describes how the weights change across training sessions. We assume that w_H returns to a baseline value of zero after each training period, and that v perfectly integrates. During a training session $k \geq 0$, the model learns a fraction q of the change in w_H that would fully minimize the error. From Eqs. (8) and (11), this can be written as

$$w_H^{(k)} = -\frac{q}{k_E k_{\text{PF}} w_{\text{PC}}} \left(\hat{g}^{(k)} - k_E k_{\text{MF}} v^{(k-1)} \right), \tag{29}$$

so that the output of the circuit after training has measured gain

$$g_{\text{train}}^{(k)} = q\hat{g}^{(k)} + (1-q) k_E k_{\text{MF}} v^{(k-1)}, \tag{30}$$

where we assume that the change in v during training is negligible, and that $\hat{g}^{(k)} = w_H^{(k)} = v^{(k)} = 0$ for all $k < 0$. During the subsequent post-training period, the model consolidates a fraction p_w of the learned weight change in w_H into v,

$$v^{(k)} = v^{(k-1)} - p_w w_H^{(k)}, \tag{31}$$

and w_H returns to a baseline value of zero, so the measured gain after consolidation is

$$g_{\text{cons}}^{(k)} = k_E k_{\text{MF}} v^{(k)}. \tag{32}$$

Then, the relative gain change consolidated is

$$p = \frac{g_{\text{cons}}^{(k)} - g_{\text{cons}}^{(k-1)}}{g_{\text{train}}^{(k)} - g_{\text{cons}}^{(k-1)}} = \frac{p_w k_{\text{MF}}}{k_{\text{PF}} w_{\text{PC}}}, \tag{33}$$

and we can rewrite the change in v after consolidation and the measured gain as

$$v^{(k)} = (1-pq)v^{(k-1)} + pq\left(\frac{\hat{g}^{(k)}}{k_{\mathrm{MF}}k_E}\right) \tag{34}$$

$$g_{\mathrm{cons}}^{(k)} = (1-pq)g_{\mathrm{cons}}^{(k-1)} + pq\hat{g}^{(k)}. \tag{35}$$

To solve the recurrence relation above, we take the unilateral z-transform of Eq. (35):

$$G_{\mathrm{cons}}^{(z)} = (1-pq)z^{-1}G_{\mathrm{cons}}^{(z)} + pq\hat{G}^{(z)},$$

where the capital letters represent z-transformed functions. We can rearrange this as

$$G_{\mathrm{cons}}^{(z)} = \frac{pq}{1-(1-pq)z^{-1}}\hat{G}^{(z)},$$

and take the inverse z-transform to find

$$g_{\mathrm{cons}}^{(k)} = \left((1-pq)^k u(k) * \hat{g}^{(k)}\right) = pq\left(e^{k\log(1-pq)}u(k) * \hat{g}^{(k)}\right), \tag{36}$$

where the star operator represents discrete convolution and the discrete unit step function is defined to be $u(k) = 1$ for $k \geq 0$ and 0 for $k < 0$. Intuitively, we can understand the measured gain after consolidation as taking a leaky average of the noisy target gain signal $\hat{g}^{(k)}$ over the recent past with time constant (in terms of training sessions) approximately $-1/\log(1-pq)$.

Given that $\hat{g}^{(k)}$ is a random variable, we can also understand Eq. (35) as an autoregressive AR(1) process, whose stationary distribution has mean and distribution

$$\mathbb{E}[g_{\mathrm{cons}}^{(\infty)}] = \mathbb{E}[\hat{g}^{(k)}] \tag{37}$$

$$\mathrm{Var}(g_{\mathrm{cons}}^{(\infty)}) = \frac{(pq)^2}{1-(1-pq)^2}\mathrm{Var}(\hat{g}^{(k)}). \tag{38}$$

That is, the stationary distribution of the measured gain after consolidation has mean equal to the true mean of the noisy target gain distribution, and variance that is an increasing function of the consolidation fraction p.

From Eq. (36) we can see that a system that consolidates a large fraction of what is learned during the training period, $p \to 1$, averages over a relatively small amount of the recent past—for fast early-learning, $q = 1$, the time constant approaches zero as $p \to 1$. Furthermore, if the mean of the noisy target gain distribution changes, the approach of the consolidated gain to the new true mean will be fast. However, as a result, the variance of the measured gain across training sessions will have a relatively large variance, from Eq. (38). On the other hand, a system that consolidates slowly ($p \to 0$) averages over a longer period, reaching the true mean slowly but with a lower stationary variance.

These features of the model are illustrated in Figure 4*F*. We plotted the value of the consolidated gain $g_{\text{cons}}^{(k)}$ resulting from simulating the simplified model, Eqs. (29) and (31), with $p = 0.1$ ("slow consolidation", magenta) or $p = 0.75$ ("fast consolidation", cyan), in response to the same set of target gain values. Initial values of w_H and v and all other parameters are the same as for the full model above (Table 1). For the first 50 training sessions, target gain values were drawn from a distribution with mean 0.4 and standard deviation 0.1, and for the subsequent 150 sessions, from a distribution with the same standard deviation but with mean 2.

Comparing a one-site and two-site model To see the advantage of the averaging effect described above, we consider a one site version of the simplified model (effectively, the late-learning weight is constant), with weight $w_{H,1}$ and no consolidation. From Eq. (8), the measured gain of this circuit after each training period k is then given by

$$g_1^{(k)} = k_E(k_{\text{MF}}v - k_{\text{PF}}w_{\text{PC}}w_{H,1}^{(k)}),$$

where v is fixed at the initial value v_0. During each training period, the model again learns a fraction q of the total weight change in $w_{H,1}$ required to minimize error,

$$\begin{aligned} w_{H,1}^{(k)} &= w_{H,1}^{(k-1)} + q(w_1^* - w_{H,1}^{(k-1)}) \\ &= (1-q)w_{H,1}^{(k-1)} - \frac{q}{k_E k_{\text{PF}} w_{\text{PC}}}\left(\hat{g}^{(k)} - k_E k_{\text{MF}} v\right), \end{aligned} \tag{39}$$

or equivalently in terms of the measured gain,

$$g_1^{(k)} = (1-q)g_1^{(k-1)} + q\hat{g}^{(k)}. \tag{40}$$

Note that the weight update, Eq. (39), is equivalent to the weight update for the two-site model, Eq. (29), if the weight were reset post-training, i.e., if $w_{H,1}^{(k-1)} = 0$ after consolidation. That is, just as for the two site model, the measured gain of the system is an AR(1) process representing a leaky average of the target gain signal with time constant $-1/\log(1-q)$. For a sequence of target gain values drawn from a normal distribution, the stationary distribution of the measured gain of the circuit will have mean equal to the true mean and variance

$$\text{Var}(g_1^{(\infty)}) = \frac{q^2}{1-(1-q)^2}\text{Var}(\hat{g}^{(k)}) \tag{41}$$

which is minimized as q becomes small ("w_H slow" in Fig. 4*A–C*). For such a small value of q, the measured gain will remain relatively stable across training sessions, but the circuit will be unable to quickly correct errors. That is, after each training session, the mean squared error of the one-site

model in the stationary limit will be

$$\begin{aligned}\mathrm{MSE}_{\mathrm{train},1} &= \lim_{k\to\infty} \mathbb{E}[(\hat{g}^{(k)} - g_1^{(k)})^2] \\ &= (1-q)^2 \lim_{k\to\infty} \mathbb{E}[(\hat{g}^{(k)} - \hat{g}_1^{(k-1)})^2] \\ &= (1-q)^2 \mathrm{MSE}_{\mathrm{pre},1}, \end{aligned} \tag{42}$$

where $\mathrm{MSE}_{\mathrm{pre},1}$ is the mean squared error at the start of each training session,

$$\begin{aligned}\mathrm{MSE}_{\mathrm{pre},1} &= \lim_{k\to\infty} \mathbb{E}[(\hat{g}^{(k)} - g_1^{(k-1)})^2] \\ &= \mathbb{E}[(\hat{g}^{(k)})^2] - \lim_{k\to\infty} \left(2\mathbb{E}[\hat{g}^{(k)}]\mathbb{E}[g_1^{(k-1)}] - \mathbb{E}[(g_1^{(k-1)})^2] \right) \\ &= \frac{2}{2-q} \mathrm{Var}(\hat{g}^{(k)}). \end{aligned} \tag{43}$$

That is, after training to adapt to a target gain experienced during training session $k-1$, $\mathrm{MSE}_{\mathrm{pre},1}$ represents the mean squared error expected at the start of training session k, when a new target gain value will be drawn. From Eqs. (42) and (43), we see that the average error after training is a decreasing function of q, whereas the average error at the start of training is an increasing function of q (for $q < 2$). Hence, there is a tradeoff between the ability of the circuit to correct errors quickly during a training session and the size of the expected future error given a noisy target gain.

Repeating this calculation for the model with two sites of learning, we have

$$\begin{aligned}\mathrm{MSE}_{\mathrm{train},2} &= \lim_{k\to\infty} \mathbb{E}[(\hat{g}^{(k)} - g_{\mathrm{train}}^{(k)})^2] \\ &= (1-q)^2 \lim_{k\to\infty} \mathbb{E}[(\hat{g}^{(k)} - g_{\mathrm{cons}}^{(k-1)})^2] \\ &= (1-q)^2 \mathrm{MSE}_{\mathrm{pre},2}, \end{aligned} \tag{44}$$

and

$$\begin{aligned}\mathrm{MSE}_{\mathrm{pre},2} &= \lim_{k\to\infty} \mathbb{E}[(\hat{g}^{(k)} - g_{\mathrm{cons}}^{(k-1)})^2] \\ &= \frac{2}{2-pq} \mathrm{Var}(\hat{g}^{(k)}). \end{aligned} \tag{45}$$

Thus, the average error after training can be reduced by fast learning at the early-learning site ($q \to 1$), and the average error at the start of training can be simultaneously reduced by slow learning at the late-learning site ($p \to 0$).

In Figure 4*C,E*, we plotted the one-site tradeoff curve as functions of the early site learning rate q ("w_H slow" and "w_H fast" arrows), with $\mathrm{MSE}_{\mathrm{pre},1}$ on the x-axis and $\mathrm{MSE}_{\mathrm{train},1}$ on the y-axis (black line in Fig. 4*C* and grey line in *E*), normalized by plotting in units of the variance of the target gain distribution. For the two-site model (Fig. 4*E*), we plotted the normalized $\mathrm{MSE}_{\mathrm{pre},2}$ on the x-axis and normalized $\mathrm{MSE}_{\mathrm{train},2}$ on the y-axis for $p = 0.1$ ("v slow", magenta) and $p = 0.75$ ("v fast",

cyan), while again varying q. In Figure 4*B*, we plotted the post-training gain $g_1^{(k)}$ from a simulation of the simple one-site model with $q = 0.1$ ("slow learning", black) or $q = 0.75$ ("fast learning", grey) over 200 training sessions, where the target gain value for each session was independently drawn from a normal distribution with mean 0.4 and standard deviation 0.1. The values of $\mathrm{MSE}_{\mathrm{pre},1}$ and $\mathrm{MSE}_{\mathrm{train},1}$ corresponding to these values of q, calculated from Eqs. (43) and (42) are plotted in Figure 4*C* with circles.

To build intuition for the difference between the one-site and two-site models more concretely, in Figure 4*A,D* we show the time course of the gain and the corresponding squared error (retinal slip) for two subsequent example training sessions. Figure 4*D* was generated by simulating the full two-site model with weight dynamics controlled by Eqs. (19) and (14). We set the learning rate at v, $k_{v,\mathrm{hetero}}$, to the value corresponding to $p = 0.1$ using Eq. (27). All other parameters were the same as those in Table 1, and the initial value of v was set so that the system would have an initial gain of 0.4, as above. To similarly simulate the time course of learning in the one-site model (Fig. 4*A*), we modified the full two-site model by removing learning at v, and the learning at w_H was governed by

$$\frac{\mathrm{d}w_H}{\mathrm{d}t} = k_w(k_{\mathrm{LTP}}\langle \mathrm{PF}(t)\rangle_{\tau_{f,w}} - k_{\mathrm{LTD}}\langle \mathrm{PF}(t)\mathrm{CF}(t)\rangle_{\tau_{f,w}}), \tag{46}$$

where we set the learning rate parameter $k_w = 0.191\ \mathrm{h}^{-1}$ (light colors) or $k_w = 1.97\ \mathrm{h}^{-1}$ (dark colors), so that for the two training sessions shown the fraction learned was $q \approx 0.10$ for the kth session and $q \approx 0.11$ for the $(k+1)$th session (light colors), or $q \approx 0.75$ for the kth and $q \approx 0.80$ for the $(k+1)$th sessions (dark colors). Unlike the two-site model, we set $k_{\mathrm{LTP}} = 0.648$ s/sp, so that outside of the training period, $\mathrm{d}w/\mathrm{d}t \approx 0$. The model used the same initial values, as well as all other parameters not specified here, as the two-site model for w_H and v. For both models, we simulated two periods of 0.5 h training followed by 11.5 h post-training without feedback about behavioral errors (post-training periods clipped for illustrative purposes).

Circuit model with internal feedback loop

Here we describe the expanded version of the circuit model that includes both a feedforward pathway carrying the sensory (vestibular) input to the early-learning area (Purkinje cells), as well as an internal feedback loop carrying the output from the late-learning area (MVN) back to the early-learning area (Fig. 6*A,F*). In the context of oculomotor learning, this internal feedback loop represents an efference copy of the output motor command. This expanded model has the same architecture as the Lisberger-Sejnowski model of oculomotor learning (Lisberger & Sejnowski 1992), but here we model the dynamics of the granule cell pathways as instantaneous. We show two candidate circuit mechanisms by which stable consolidation can be achieved when internal feedback is present.

The vestibular stimulus is carried by mossy fibers, described as before by Eq. (1), and by

vestibular parallel fibers with firing rate

$$\mathrm{PF}_H(t) = \mathrm{PF}_0 + k_{\mathrm{PF},H}\dot{H}(t). \tag{47}$$

The efference copy is carried through a different set of parallel fibers with firing rate

$$\mathrm{PF}_E(t) = \mathrm{PF}_0 + k_{\mathrm{PF},E}\dot{E}(t), \tag{48}$$

where the eye velocity output, $\dot{E}(t)$, is defined as before in Eq. (5). The firing rate of Purkinje cells is then a linear combination of vestibular and efference copy input, with a spontaneous baseline rate, i.e.,

$$\mathrm{PC}(t) = \mathrm{PC}_0 + w_H(t)\mathrm{PF}_H(t) + w_E(t)\mathrm{PF}_E(t). \tag{49}$$

We assume that both parallel fiber pathways result in monosynaptic excitation and disynaptic inhibition onto the Purkinje cell, so that

$$w_H(t) = w_H^+(t) - w_H^-, \tag{50}$$

$$w_E(t) = w_E^+(t) - w_E^-, \tag{51}$$

where w_H^+ and w_E^+ are the excitatory weights of the feedforward and feedback parallel fibers and w_H^- and w_E^- are the inhibitory weights of the interneurons in each pathway. Thus, as before, we can interpret $w_H = 0$ and $w_E = 0$ as balanced excitatory and inhibitory weights for each pathway. The firing rate of MVN neurons is defined as before in the first line of Eq. (4), except we substitute the new definition of Purkinje cell activity from Eq. (49). The late-learning weight of the direct pathway is still called v. Then, following the same reasoning as the derivation leading to Eq. (8), we can calculate the eye velocity output as

$$\dot{E}(t) \approx -\frac{k_E(\tilde{v}(t) - w_{\mathrm{PC}}\tilde{w}_H(t))}{1 - w_{\mathrm{PC}}k_E\tilde{w}_E(t)}\dot{H}(t) = -g(t)\dot{H}(t), \tag{52}$$

defining the prefactor in the second equality as the gain $g(t)$, and where for simplicity we let

$$\tilde{v}(t) = k_{\mathrm{MF}} \cdot v(t), \tag{53}$$

$$\tilde{w}_H(t) = k_{\mathrm{PF},H} \cdot w_H(t), \tag{54}$$

$$\text{and } \tilde{w}_E(t) = k_{\mathrm{PF},E} \cdot w_E(t). \tag{55}$$

We use modified learning rules for the excitatory weights in the parallel fiber pathways, in which

the weight values do not passively decay,

$$\begin{aligned}\frac{\mathrm{d}w_x^+}{\mathrm{d}t} &= k_{\mathrm{LTP}}\langle \mathrm{PF}_x(t)\rangle_{\tau_{f,w}} - k_{\mathrm{LTD}}\langle \mathrm{PF}_x(t)\cdot \mathrm{CF}(t)\rangle_{\tau_{f,w}} \\ &\approx w_{x,\infty}^+ - c_x(t),\end{aligned} \tag{56}$$

where $x = H$ or $x = E$ for the head pathway or efference copy pathway respectively, and $w_{H,\infty}^+ = w_{E,\infty}^+ = \mathrm{PF}_0(k_{\mathrm{LTP}} - k_{\mathrm{LTD}}\mathrm{CF}_0)$ and $c_x(t) = k_{\mathrm{LTD}}\langle \delta\mathrm{PF}_x(t)\delta\mathrm{CF}(t)\rangle_{\tau_{f,w}}$. In the following analysis, we will assume that w_E has a fixed nonnegative value for the purpose of visualizing the dynamics, but stable consolidation was also possible when w_E was plastic (Fig. S3; SI Text, S6). So that plasticity is completely controlled by the correlation between parallel fibers and climbing fibers, we assumed values of k_{LTD} and k_{LTP} such that $w_{H,\infty}^+ = 0$. This condition can be interpreted as the plasticity due to spontaneous PF activity being offset by spontaneous CF activity (Kenyon *et al.* 1998).

For the direct pathway weight v, we used a modified version of the heterosynaptic rule in Eq. (14),

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -k_{v,\mathrm{hetero}}\langle \mathrm{MF}(t)(\mathrm{PC}(t) - \langle \mathrm{PC}(t)\rangle_{\tau_f})\rangle_{\tau_{f,v}}, \tag{57}$$

where the subtraction of the fixed Purkinje cell baseline has been replaced with a sliding threshold equal to the average PC activity over the recent past, similar to that used in the covariance-like Hebb rule, Eq. (15). The time constant of the sliding threshold was τ_f, which was slow relative to variations in the sensory input but fast compared to the rates of plasticity (see Eq. (6) and SI Text, S5 for how this affects stability). Following the same line of reasoning leading to Eq. (20), we can rewrite the heterosynaptic plasticity rule in terms of the weights as

$$\frac{\mathrm{d}v}{\mathrm{d}t} \approx -k_{v,\mathrm{hetero}}\frac{k_{\mathrm{MF}}(\tilde{w}_H(t) - k_E\tilde{w}_E\tilde{v}(t))}{1 - w_{\mathrm{PC}}k_E\tilde{w}_E}\sigma_H^2. \tag{58}$$

Note that, unlike the feedforward model, Eq. (58) implies that there must be variability in the post-training input, $\sigma_H^2 > 0$, for consolidation to occur.

For stable consolidation, we need the learning rule for the late-learning site to stably hold any value of v. With the heterosynaptic plasticity rule, from Eq. (58) this occurs as long as the feedforward weight at the early-learning weight is

$$w_H = \frac{k_E}{k_{\mathrm{PF},H}}\tilde{w}_E\tilde{v}, \tag{59}$$

which defines a line in w_H-v space (Fig. 6*E,J*). Values of w_H and v that lie along this line cause the early-learning area to not respond to sensory input to the circuit, i.e., the reset condition is satisfied. If $w_E > 0$, this can no longer be achieved in general by a passive decay at w_H to zero, as w_H needs to go to a positive steady state value since v is also positive. Here we propose two potential circuit mechanisms, each of which can properly reset the early-learning activity and thus stably consolidate learning.

Dynamics during post-training period

Reset Mechanism 1: Plasticity of inhibitory input to early-learning area First, stable consolidation can be achieved if the feedforward inhibitory synapses onto the early-learning area (w_H^-) are plastic, governed by a Hebbian covariance-like rule,

$$\frac{\mathrm{d}w_H^-}{\mathrm{d}t} = k_{\mathrm{inh}}\langle \mathrm{PF}_H(t)(\mathrm{PC}(t) - \langle \mathrm{PC}(t)\rangle_{\tau_f})\rangle_{\tau_{\mathrm{inh}}}. \tag{60}$$

During the post-training period, the correlation between climbing fibers and the feedforward parallel fibers is zero, $c_H(t) \equiv 0$, so simplifying as for Eq. (58), and combining with the learning rule for the excitatory weight, Eq. (56), we can write the equation for the overall feedforward weight at the early-learning site as

$$\frac{\mathrm{d}w_H}{\mathrm{d}t} \approx -\frac{k_{\mathrm{inh}}k_{\mathrm{PF},H}^2\sigma_H^2}{1 - w_{\mathrm{PC}}k_E\tilde{w}_E}\left(w_H - \frac{k_E}{k_{\mathrm{PF},H}}\tilde{w}_E\tilde{v}\right), \tag{61}$$

which has a fixed point that lies along the line in Eq. (59), satisfying the reset condition, as long as

$$w_E < \frac{1}{k_{\mathrm{PF},E}w_{\mathrm{PC}}k_E}. \tag{62}$$

To determine whether the fixed point is attractive, it suffices to show that the slope of the trajectory approaching the fixed point is smaller than the slope of the line in Eq. (59). Otherwise, trajectories will move continuously away from the line of fixed points. From Eqs. (61) and (58), trajectories have slope

$$\frac{\mathrm{d}v}{\mathrm{d}w_H} \approx \frac{k_{v,\mathrm{hetero}}k_{\mathrm{MF}}}{k_{\mathrm{inh}}k_{\mathrm{PF},H}}, \tag{63}$$

whereas from Eq. (59) the slope of the line of fixed points is $k_{\mathrm{PF},H}/(k_E\tilde{w}_E k_{\mathrm{MF}})$. Thus, the slope of trajectories is smaller if

$$\frac{k_{v,\mathrm{hetero}}k_{\mathrm{MF}}^2k_E\tilde{w}_E}{k_{\mathrm{inh}}k_{\mathrm{PF},H}^2} = \frac{\tau_{w_H,\mathrm{dark}}(\sigma_H^2)}{\tau_v(\sigma_H^2)} < 1, \tag{64}$$

where the time constants $\tau_{w_H,\mathrm{dark}}$ and τ_v are defined from Eqs. (61) and (58) respectively as functions of the input variability σ_H^2,

$$\tau_{w_H,\mathrm{dark}}(\sigma_H^2) = \frac{1 - w_{\mathrm{PC}}k_E\tilde{w}_E}{k_{\mathrm{reset}}k_{\mathrm{PF},H}^2\sigma_H^2} \tag{65}$$

$$\tau_v(\sigma_H^2) = \frac{1 - w_{\mathrm{PC}}k_E\tilde{w}_E}{k_{v,\mathrm{hetero}}k_{\mathrm{MF}}^2k_E\tilde{w}_E\sigma_H^2}, \tag{66}$$

and where, to highlight the similarity of these expressions to the analogous quantities for the other reset mechanism below, we have defined

$$k_{\mathrm{reset}} = k_{\mathrm{inh}} \quad \text{(inhibitory plasticity reset mechanism).} \tag{67}$$

In other words, the line in Eq. (59) is an attractor only if plasticity is faster at the early-learning site than at the late-learning site.

The inhibitory plasticity here drives w_H so that the activity of the early-learning area does not modulate in response to sensory input. This result is broadly consistent with previous work showing that Hebbian plasticity of inhibitory synapses tends toward restoring excitatory-inhibitory balance (Vogels *et al.* 2011), except here this balance is only in terms of the fluctuating components of the inputs and not the spontaneous components, because of the sliding threshold term in Eq. (60). Note that, if $\sigma_H^2 = 0$, early-learning activity will not be reset, but there will also be no consolidation at v.

Reset Mechanism 2: Inhibition of pathway carrying instructive signals Alternatively, stable consolidation can be achieved if, in addition to providing information about behavioral errors, the climbing fiber input at the early-learning site also serves to reset early-learning activity. We modeled climbing fiber activity as being driven by the instructive signal and inhibited by medial vestibular nucleus cells that receive only Purkinje cell inhibition (Fig. 6*F*), which corresponds to net excitation by PC activity,

$$\delta\mathrm{CF}(t) = k_{\mathrm{CF}} \tanh\left(-\beta_{\mathrm{light}}\dot{R}(t) + \beta_{\mathrm{reset}}\delta\mathrm{PC}(t)\right), \tag{68}$$

where $\delta\mathrm{PC}(t) = \mathrm{PC}(t) - \langle \mathrm{PC}(t)\rangle_{\tau_f}$. When information about errors is not present, $\dot{R} \equiv 0$, so that for small errors, the parallel fiber-climbing fiber covariance $c_H(t)$ is in the linear regime, and the learning rule for w_H becomes

$$\frac{\mathrm{d}w_H}{\mathrm{d}t} \approx -\frac{k_{\mathrm{LTD}}k_{\mathrm{CF}}\beta_{\mathrm{reset}}k_{\mathrm{PF},H}^2\sigma_H^2}{1 - w_{\mathrm{PC}}k_E\tilde{w}_E}\left(w_H - \frac{k_E}{k_{\mathrm{PF},H}}\tilde{w}_E\tilde{v}\right), \tag{69}$$

which has the same fixed points as Eq. (61) that satisfy the reset condition if Eq. (62) holds. Similarly, as in Eq. (64), the fixed point for a given trajectory is attractive if $\tau_{w_H,\mathrm{dark}} < \tau_v$, where $\tau_{w_H,\mathrm{dark}}$ and τ_v are defined as in Eqs. (65) and (66) but where we now define

$$k_{\mathrm{reset}} = k_{\mathrm{LTD}}k_{\mathrm{CF}}\beta_{\mathrm{reset}} \quad \text{(inhibition of instructive signals reset mechanism).} \tag{70}$$

For large errors, the magnitude of $\delta\mathrm{CF}$ will be saturated, and $\mathrm{d}w_H/\mathrm{d}t$ will reach a maximum value, so the slope of trajectories will be

$$\frac{\mathrm{d}v}{\mathrm{d}w_H} \approx \frac{k_{v,\mathrm{hetero}}k_{\mathrm{MF}}\sigma_H^2}{k_{\mathrm{LTD}}k_{\mathrm{CF}}C(1 - w_{\mathrm{PC}}k_E\tilde{w}_E)}\left(w_H - \frac{k_E}{k_{\mathrm{PF},H}}\tilde{w}_E\tilde{v}\right), \tag{71}$$

where C is defined as in Eq. (22). That is, the trajectories become increasingly vertical the further they are from the line of fixed points. For trajectories to converge to a stable fixed point, they

must have slopes shallower than the line of fixed points. This occurs when

$$\left| w_H - \frac{k_E}{k_{\mathrm{PF},H}} \tilde{w}_E \tilde{v} \right| < \frac{k_{\mathrm{LTD}} k_{\mathrm{CF}} k_{\mathrm{PF},H} C (1 - w_{\mathrm{PC}} k_E \tilde{w}_E)}{k_{v,\mathrm{hetero}} k_{\mathrm{MF}}^2 \sigma_H^2 k_E \tilde{w}_E}. \tag{72}$$

This effectively places a limit on how large a gain change the circuit can learn during the training period in order for that gain change to be stably consolidated (see SI Text, S5).

Dynamics during training

During the training period, plasticity at the early-learning site is also driven by information about behavioral errors through the retinal slip error $\dot{R}$, as defined in Eq. (11). For either candidate post-training reset mechanism, using information about errors, the circuit can stably learn a graded change in gain if the effective rate of plasticity of w_H is faster than that of v (Fig. S4; SI Text, S5). This condition is stricter than Eq. (64) and, if met, automatically guarantees stability during the post-training period. For the inhibition of instructive signals mechanism specifically, there is a maximum gain change that can be learned during a training period, as alluded to in the previous section.

Effect of post-training reset on weight dynamics

With either the inhibitory plasticity or homeostatic climbing fiber mechanism, two important properties arise in the model's dynamics that are also observed experimentally (see SI Text, S5). First, the early-learning weight w_H tends toward a saturation point during training that is not simply the value that would minimize errors. In the feedforward model, this property was modeled explicitly by including a decay term in the learning rule Eq. (13) (following previous work), which we do not include here (i.e., in Eq. (56)). Second, during training, the approach of the early-learning weight to the saturation point occurs over a faster timescale than the subsequent post-training reset during consolidation (for the inhibition of instructive signals, this is true in the linear regime of the CF response). This was also modeled explicitly in the feedforward model with two different values of τ_w during the training and post-training periods, but here results from the fact that learning and resetting are being driven by two different mechanisms. Note that we could also include either of the reset mechanisms in the feedforward model (as it is the case of $w_E = 0$), and the resulting dynamics would also have these properties without having to explicitly model them.

Simulation of oculomotor learning in the model with feedback

We simulated the same oculomotor experiment as for the model without feedback (see *Simulation of oculomotor learning*), with differences in parameters specified in Tables 2 and 3. We used the same sensory input as for the feedforward model during both the training and post-training periods. We simulated the model with both the inhibitory plasticity (Fig. 6*A*–*E*) and inhibition of instructive

signals (Fig. 6*F*–*J*) reset mechanism, assuming that w_E is not plastic. We used a target gain value of $g^{\text{target}} = 2$, and initial values $w^+_{H,0}$ and v_0 for w^+_H and v, respectively.

Plasticity of internal feedback weight

Finally, we built a version of the model in which the strength of the feedback weight w_E was plastic (SI Text, S6). We found that stable learning and consolidation could still be achieved using either of the two reset mechanisms if plasticity at the early-learning sites was fast compared to plasticity at the late-learning site, but that the weight w_E needed to be kept within a particular range for stability (SI Text, S6). In Figure S3, we simulated the model with the inhibition of instructives signal mechanism. All parameters were as for the non-plastic feedback model (Tables 1–3).

Table 1: Values and descriptions for parameters of feedforward circuit model. Sources for values taken from the literature are provided in the right-hand column. Parameters that were calculated based on published values are indicated as "based on" a source.

Parameter	Value	Description	Source
MF_0	55 sp/s	Mossy fiber (MF) spontaneous baseline rate	Lasker *et al.* (2008)
k_{MF}	0.14 (sp/s)/(deg/s)	MF sensitivity to vestibular input	
PF_0	14 sp/s	Parallel fiber (PF) spontaneous baseline rate	Based on Arenz *et al.* (2008)
k_{PF}	0.42 (sp/s)/(deg/s)	PF sensitivity to vestibular input	
PC_0	50 sp/s	Purkinje cell spontaneous baseline rate	Katoh *et al.* (2015)
MVN_0	-12 sp/s	Correction factor to make medial vestibular nucleus (MVN) spontaneous baseline rate = 57 sp/s	Beraneck & Cullen (2007)
CF_0	1 sp/s	Climbing fiber (CF) spontaneous baseline rate	E.g., Goossens *et al.* (2004) and Maruta *et al.* (2007)
k_{CF}	1 sp/s	Maximum amplitude of CF response around baseline	
β_{light}	1 s/deg	Scale factor for saturation of CF response	
k_E	2.2 (deg/s)/(sp/s)	Eye velocity sensitivity to MVN firing	Based on Beraneck & Cullen (2007)
k_{LTP}	1.005 s/sp	Contribution of parallel fiber firing to plasticity at early-learning site w_H	Based on amount of learning in Kimpo *et al.* (2014) and Boyden & Raymond (2003)
k_{LTD}	0.648 $(\mathrm{s/sp})^2$	Contribution of parallel fiber-climbing fiber coincidence to plasticity at w_H	
$\tau_{w,\mathrm{train}}$	0.15 h	Time constant of plasticity at w_H during training	Boyden & Raymond (2003)
$\tau_{w,\mathrm{post}}$	5 h	Time constant of plasticity at w_H after training	Based on Cooke *et al.* (2004) and Jang *et al.* (2020)
$k_{v,\mathrm{hetero}}$	2.75×10^{-5} $(\mathrm{s/sp})^2/\mathrm{h}$	Plasticity rate at late-learning site v for heterosynaptic rule	
$k_{v,\mathrm{Hebb}}$	8×10^{-3} $(\mathrm{s/sp})^2/\mathrm{h}$	Plasticity rate at v for Hebbian rule	

Continues on next page

Table 1 continued.

Parameter	Value	Description	Source
τ_f	1 min	Timescale of fast average used to calculate eye movement output	
$\tau_{f,w}$	1 min	Timescale of average used in plasticity rule at w_H	
$\tau_{f,v}$	1 min (Fig. 2) 0.7 h (Fig. 3)	Timescale of average used in plasticity rule at v	
τ_s	0.0395 h	Timescale of sliding average θ in Hebbian rule for v	
w_H^-	5	Nonplastic molecular layer interneuron to Purkinje cell weight	Consistent with weight change in Jang *et al.* (2020)
$w_{H,0}^+$	5	Initial value of early-learning, PF-Purkinje cell weight w_H	
w_{PC}	0.05	Purkinje cell to MVN synaptic strength	Based on Payne *et al.* (2019)
v_0	1.3	Initial value of late-learning MF-MVN weight v	

Table 2: Values, descriptions and sources for parameters of circuit model with internal efference copy feedback, using the climbing fiber negative feedback reset mechanism. All other parameters same as in Table 1.

Parameter	Value	Description	Source
$k_{\mathrm{PF},H}$	0.42 (sp/s)/(deg/s)	PF sensitivity to vestibular input	Based on Arenz *et al.* (2008)
$k_{\mathrm{PF},E}$	0.42 (sp/s)/(deg/s)	PF sensitivity to efference copy input	
PC_0	-75 sp/s	Correction factor to make Purkinje cell spontaneous baseline rate = 50 sp/s	Katoh *et al.* (2015)
MVN_0	-9.9 sp/s	Correction factor to make MVN spontaneous baseline = 57 sp/s	Beraneck & Cullen (2007)
k_{LTD}	0.75 $(\mathrm{s/sp})^2/\mathrm{h}$	Contribution of PF firing to plasticity at early-learning site w_H	
k_{LTP}	0.75 (s/sp)/h	Contribution of PF-CF coincidence to plasticity at w_H	
β_{light}	1 s/deg	Slope parameter for error-driven plasticity	
β_{reset}	0.0125 s/sp	Slope parameter for reset plasticity	
$k_{v,\mathrm{hetero}}$	3.6×10^{-3} $(\mathrm{s/sp})^2/\mathrm{h}$	Plasticity rate at v	
$w^+_{H,0}$	2.5	Initial value of excitatory feedforward early-learning weight w^+_H	
w^-_H	5	Nonplastic molecular layer interneuron to Purkinje cell weight	Consistent with weight change in Jang *et al.* (2020)
w^+_E	8.43	Nonplastic excitatory weight of efference copy input to Purkinje cell	Net weight based on Katoh *et al.* (2015)
w^-_E	2	Nonplastic inhibitory weight of efference copy input to Purkinje cell	
v_0	1.26	Initial value of late-learning weight	

Table 3: Values and descriptions of parameters of circuit model with internal efference copy feedback, using the inhibitory plasticity reset mechanism. All other parameters same as in Tables 1 and 2.

Parameter	Value	Description
k_{inh}	0.01 $(\mathrm{s/sp})^2/\mathrm{h}$	Maximum rate of reset plasticity
$k_{v,\mathrm{hetero}}$	3.9×10^{-3} $(\mathrm{s/sp})^2/\mathrm{h}$	Plasticity rate at v
τ_{inh}	1 min	Timescale of sliding average of PC activity

References

107. Lasker, D. M., Han, G. C., Park, H. J. & Minor, L. B. Rotational Responses of Vestibular–Nerve Afferents Innervating the Semicircular Canals in the C57BL/6 Mouse. *Journal of the Association for Research in Otolaryngology* **9,** 334–348. doi:10.1007/s10162-008-0120-4 (2008).

108. Arenz, A., Silver, R. A., Schaefer, A. T. & Margrie, T. W. The Contribution of Single Synapses to Sensory Representation in Vivo. *Science* **321,** 977–980. doi:10.1126/science.1158391 (2008).

109. Katoh, A., Shin, S.-L., Kimpo, R. R., Rinaldi, J. M. & Raymond, J. L. Purkinje Cell Responses during Visually and Vestibularly Driven Smooth Eye Movements in Mice. *Brain and Behavior* **5,** e00310. doi:10.1002/brb3.310 (2015).

110. Beraneck, M. & Cullen, K. E. Activity of Vestibular Nuclei Neurons During Vestibular and Optokinetic Stimulation in the Alert Mouse. *Journal of Neurophysiology* **98,** 1549–1565. doi:10.1152/jn.00590.2007 (2007).

111. Goossens, H. H. L. M., Hoebeek, F. E., van Alphen, A. M., van der Steen, J., Stahl, J. S., De Zeeuw, C. I. & Frens, M. A. Simple Spike and Complex Spike Activity of Floccular Purkinje Cells during the Optokinetic Reflex in Mice Lacking Cerebellar Long-Term Depression. *European Journal of Neuroscience* **19,** 687–697. doi:10.1111/j.0953-816X.2003.03173.x (2004).

112. Maruta, J., Hensbroek, R. A. & Simpson, J. I. Intraburst and Interburst Signaling by Climbing Fibers. *Journal of Neuroscience* **27,** 11263–11270. doi:10.1523/JNEUROSCI.2559-07.2007 (2007).

113. Boyden, E. S. & Raymond, J. L. Active Reversal of Motor Memories Reveals Rules Governing Memory Encoding. *Neuron* **39,** 1031–1042. doi:10.1016/s0896-6273(03)00562-2 (2003).

114. Payne, H. L., French, R. L., Guo, C. C., Nguyen-Vu, T. B., Manninen, T. & Raymond, J. L. Cerebellar Purkinje Cells Control Eye Movements with a Rapid Rate Code That Is Invariant to Spike Irregularity. *eLife* **8,** e37102. doi:10.7554/eLife.37102 (2019).

115. Bagnall, M. W., McElvain, L. E., Faulstich, M. & du Lac, S. Frequency-Independent Synaptic Transmission Supports a Linear Vestibular Behavior. *Neuron* **60,** 343–352. doi:10.1016/j.neuron.2008.10.002 (2008).

116. Guo, C. C., Ke, M. C. & Raymond, J. L. Cerebellar Encoding of Multiple Candidate Error Cues in the Service of Motor Learning. *Journal of Neuroscience* **34,** 9880–9890. doi:10.1523/JNEUROSCI.5114-13.2014 (2014).

117. Raymond, J. L. & Lisberger, S. G. Neural Learning Rules for the Vestibulo-Ocular Reflex. *Journal of Neuroscience* **18,** 9112–9129. doi:10.1523/JNEUROSCI.18-21-09112.1998 (1998).

118. Suvrathan, A., Payne, H. L. & Raymond, J. L. Timing Rules for Synaptic Plasticity Matched to Behavioral Function. *Neuron* **92,** 959–967. doi:10.1016/j.neuron.2016.10.022 (2016).

119. Lev-Ram, V., Wong, S. T., Storm, D. R. & Tsien, R. Y. A New Form of Cerebellar Long-Term Potentiation Is Postsynaptic and Depends on Nitric Oxide but Not cAMP. *Proceedings of the National Academy of Sciences of the United States of America* **99,** 8389–8393. doi:10.1073/pnas.122206399 (2002).

120. Zenke, F., Gerstner, W. & Ganguli, S. The Temporal Paradox of Hebbian Learning and Homeostatic Plasticity. *Current Opinion in Neurobiology* **43,** 166–176. doi:10.1016/j.conb.2017.03.015 (2017).

121. Zenke, F., Hennequin, G. & Gerstner, W. Synaptic Plasticity in Neural Networks Needs Homeostasis with a Fast Rate Detector. *PLoS Computational Biology* **9,** e1003330. doi:10.1371/journal.pcbi.1003330 (2013).

122. Nguyen-Vu, T. D. B., Kimpo, R. R., Rinaldi, J. M., Kohli, A., Zeng, H., Deisseroth, K. & Raymond, J. L. Cerebellar Purkinje Cell Activity Drives Motor Learning. *Nature Neuroscience* **16,** 1734–1736. doi:10.1038/nn.3576 (2013).

123. Kenyon, G. T., Medina, J. F. & Mauk, M. D. A Mathematical Model of the Cerebellar-Olivary System I: Self-Regulating Equilibrium of Climbing Fiber Activity. *Journal of Computational Neuroscience* **5,** 17–33 (1998).

SI Text

S1 Consolidation with the Hebbian rule is not robust to noisy input

The failure of the stabilized Hebbian rule, Eq. (15), to successfully consolidate memory from the early- to late-learning site in the presence of variability in the sensory input (Fig. 5) can be understood through an analysis similar to the one performed for the heterosynaptic rule (see Materials and methods, *Analysis of synaptic learning dynamics*). We again assume that the mean of the sensory input is approximately zero and that the variability σ_H^2 is constant over the timescales $\tau_{f,v}$ and τ_s. We also assume, as before, that the timescale over which the weights change is much slower than $\tau_{f,v}$. Using that the threshold $\theta(t)$ is the exponential average of $\mathrm{MVN}(t)$, we have

$$\theta(t) = \langle \mathrm{MVN}(t) \rangle_{\tau_s} \approx \mathrm{MF}_0 \langle v \rangle_{\tau_s} - w_{\mathrm{PC}} \mathrm{PF}_0 \langle w \rangle_{\tau_s} - w_{\mathrm{PC}} \mathrm{PC}_0 + \mathrm{MVN}_0. \tag{S.1}$$

Then, substituting this and the definition of $\mathrm{MVN}(t)$, Eq. (4), into the learning rule, Eq. (15) becomes

$$\frac{\mathrm{d}v}{\mathrm{d}t} \approx k_{v,\mathrm{Hebb}}[\langle \mathrm{MF}(t)^2 \rangle_{\tau_{f,v}} v - w_{\mathrm{PC}} \langle \mathrm{MF}(t)\mathrm{PF}(t) \rangle_{\tau_{f,v}} w_H - \mathrm{MF}_0^2 v_s + w_{\mathrm{PC}} \mathrm{MF}_0 \mathrm{PF}_0 w_{H,s}], \tag{S.2}$$

where $w_{H,s} = \langle w_H \rangle_{\tau_s}$ and $v_s = \langle v \rangle_{\tau_s}$.

In the post-training period, when $c \to 0$, we have that $w_H \to w_{H,0} = w_\infty^+ - w_H^-$. As w_H approaches this steady state, so will the exponential average, i.e., $w_{H,s} \to w_{H,0}$ (see Eq. (6)), so we can use this quasi-steady state assumption on w_H to simplify Eq. (S.2) further to

$$\frac{\mathrm{d}v}{\mathrm{d}t} \approx \frac{1}{\tau_v(\sigma_H^2)} v - k_{v,\mathrm{Hebb}}[w_{\mathrm{PC}} k_{\mathrm{MF}} k_{\mathrm{PF}} \sigma_H^2 w_{H,0} + \mathrm{MF}_0^2 v_s] \quad \text{for } w_H, w_{H,s} \to w_0, \tag{S.3}$$

where the time constant τ_v is a function of the variability of the sensory input, σ_H^2,

$$\tau_v(\sigma_H^2) = \frac{1}{k_{v,\mathrm{Hebb}}(\mathrm{MF}_0^2 + k_{\mathrm{MF}}^2 \sigma_H^2)}. \tag{S.4}$$

In the case that there is no variability in the sensory input, $\sigma_H^2 = 0$, we have that the dynamics of learning are given by

$$\tau_v(0) \frac{\mathrm{d}v}{\mathrm{d}t} \approx v - v_s, \tag{S.5}$$

$$\tau_s \frac{\mathrm{d}v_s}{\mathrm{d}t} = -v_s + v. \tag{S.6}$$

Here, the first eigenvalue of the system is 0, with corresponding eigenvector $[1, 1]$ so that any state

along the line $v = v_s$ is a fixed point of the system. The second eigenvalue of the dynamics is

$$\lambda = \frac{1}{\tau_v(0)} - \frac{1}{\tau_s}, \tag{S.7}$$

which is decaying if $\tau_s < \tau_v(0)$, i.e., the timescale of the sliding average is faster than the timescale of plasticity (cf. Zenke *et al.* (2013)). Thus, the synaptic learning dynamics contain a line attractor in the space of v and the internal variable v_s that does not depend on w_0. Geometrically, when $\sigma_H^2 = 0$, the nullclines for the dynamics of v and v_s overlap, creating the line attractor.

If there is variability in the sensory input, $\sigma_H^2 > 0$, the nullclines no longer exactly overlap; there is now a single fixed point $v^* = v_s^* = w_{H,0}$. This fixed point is unstable (it is a saddle point of the dynamics), with eigenvalues

$$\lambda_\pm = \left(\frac{1}{\tau_v(\sigma_H^2)} - \frac{1}{\tau_s}\right) \cdot \frac{1}{2}\left[1 \mp \sqrt{1 + \frac{4k_{v,\text{Hebb}}k_{\text{MF}}^2\sigma_H^2}{\tau_s(1/\tau_v(\sigma_H^2) - 1/\tau_s)^2}}\right],$$

implying that one of the eigenvalues will always be positive.

In principle, the instability can be rectified by adding an additional decay term $-\varepsilon v$ to the learning rule, so that Eq. (15) is

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -\varepsilon v + \langle \text{MF}(t)(\text{MVN}(t) - \theta(t))\rangle_{\tau_{f,v}}. \tag{S.8}$$

If we write the decay rate as $\varepsilon = k_{v,\text{Hebb}}k_{\text{MF}}^2\sigma^2$ in terms of a new variable σ^2, we can simplify this equation (as we did for Eq. (S.3)) to

$$\frac{\mathrm{d}v}{\mathrm{d}t} \approx k_{v,\text{Hebb}}[k_{\text{MF}}^2(\sigma_H^2 - \sigma^2)v + \text{MF}_0^2(v - v_s) - w_{\text{PC}}k_{\text{MF}}k_{\text{PF}}\sigma_H^2 w_0]. \tag{S.9}$$

If σ^2 is equal to σ_H^2, the variability of the vestibular input in some period, and if $w_{H,0} = 0$, then the system will again be described by the dynamics given by Eqs. (S.5) and (S.6). That is, $v = v_s$ will still be a line attractor. If σ^2 is not exactly equal to σ_H^2, then the weight will either grow unstably or decay to zero, depending on whether $\sigma^2 < \sigma_H^2$ or $\sigma^2 > \sigma_H^2$, respectively. This can be seen by assuming the threshold θ moves instantaneously, so that $v_s \approx v$, and noting that Eq. (S.9) becomes either exponential growth or decay, with time constant inversely proportional to the degree of mistuning. So, for successful systems consolidation in a circuit using this modified version of the Hebbian rule, the learning rule must "know" or otherwise measure the variability of the input σ_H^2. While a learning rule that accounts for the natural statistics of inputs presents an interesting possibility, we note that it may be challenging to implement biologically.

S2 Traditional homeostatic Hebbian rules do not readily support consolidation of analog memories

It has long been recognized that the basic form of the Hebbian learning rule generally leads to unstable positive feedback (Miller & MacKay 1994; Zenke *et al.* 2017; Abbott & Nelson 2000), leading to the development of rules which contain homeostatic mechanisms that attempt to counter this instability. The most well-known classes of such homeostatic mechanisms include weight normalization, in which weights are adjusted so that the weight vector for a neuron maintains a given length (von der Malsburg 1973; Oja 1982; Miller & MacKay 1994), and firing rate-target homeostasis methods, in which weights are adjusted so that the mean postsynaptic firing rate is maintained close to a specified target. The latter class of mechanisms includes those commonly called synaptic scaling (Turrigiano 2008; Yger & Gilson 2015; Tetzlaff *et al.* 2011). In order to stably consolidate analog memories, the learning rule at the late-learning site must be able to support a continuum of stable weight configurations. For the circuit model we developed here, we show below that Oja's rule and the BCM rule, prominent examples of weight normalization and firing rate-target homeostasis rules, respectively, cannot support such a stable weight continuum.

First, we examine Oja's rule, a form of weight normalization (Oja 1982). We will consider the Hebbian piece in a "covariance" form,

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \eta(\langle \delta\mathrm{MF}(t)\delta\mathrm{MVN}(t)\rangle - \langle \delta\mathrm{MVN}^2\rangle v), \tag{S.10}$$

where for simplicity, $\delta\mathrm{MF}(t)$ and $\delta\mathrm{MVN}(t)$ are baseline-subtracted forms of mossy fiber and MVN neuron firing. We assume that the timescale of the average indicated by the brackets is slow compared to variation in the input but faster than the timescale of plasticity of v (as for τ_f in the main text). Both before training and after consolidation, when $w \to 0$, so that $\mathrm{PC}(t) \to \mathrm{PC}_0$,

$$\delta\mathrm{MVN}(t) = v\delta\mathrm{MF}(t).$$

Then, the learning rule is at steady state when

$$v^*\langle \delta\mathrm{MF}^2\rangle - {v^*}^3\langle \delta\mathrm{MF}^2\rangle = 0 \Rightarrow v^* = 0 \text{ or } v^* = \pm 1.$$

This is because Oja's rule is designed to maintain the Euclidean length of the weight vector fixed at 1, and here there is only one input source with a plastic weight. Even in the case where there are multiple mossy fiber input types with plastic weights, Oja's rule will drive the weight vector so that it is parallel (or antiparallel) to the eigenvector of the covariance matrix of the inputs that has largest eigenvalue (Oja 1982). In our model, in which we do not expect the signals carried by the mossy fiber inputs to themselves change persistently as a result of training, only one stable weight configuration can be maintained. Thus, Oja's rule does not readily support consolidation in our model.

We next examine the Bienenstock-Cooper-Monroe (BCM) rule (Bienenstock *et al.* 1982), one of the most well-known firing rate-target homeostasis rules. The BCM rule consists of a Hebbian term multiplied by a term that changes the sign of plasticity depending on whether postsynaptic activity is greater than or less than a sliding threshold. For our circuit, this takes the form

$$\tau_v \frac{\mathrm{d}v}{\mathrm{d}t} = \langle \mathrm{MF}(t)\mathrm{MVN}(t)\rangle[\langle \mathrm{MVN}(t)\rangle - \theta(t)], \tag{S.11}$$

$$\tau_\theta \frac{\mathrm{d}\theta}{\mathrm{d}t} = -\theta + \frac{\langle \mathrm{MVN}(t)\rangle^2}{\mathrm{MVN}_{\mathrm{target}}}. \tag{S.12}$$

The angle brackets again represent time averages taken over a timescale longer than the variation in the input, but much shorter than the timescale of plasticity. The system has a fixed point whenever $\theta = \langle \mathrm{MVN}\rangle = \mathrm{MVN}_{\mathrm{target}}$.

There is one fixed point, with steady state weight value

$$v^* = \frac{\mathrm{MVN}_{\mathrm{target}} - \mathrm{MVN}_0 + w_{\mathrm{PC}}\langle \mathrm{PC}\rangle}{\mathrm{MF}_0}.$$

The value of v^* is determined by parameters of the system that do not change as a result of learning ($\mathrm{MF}_0, \mathrm{MVN}_{\mathrm{target}}$ and MVN_0), as well as a variable ($\langle \mathrm{PC}\rangle$) that returns to its pre-training baseline value during the post-training period. Therefore, the system with one plastic synapse cannot persistently hold all possible values of v that may be reached after training.

For multiple plastic inputs to the (linear) MVN neuron, the BCM rule is less restrictive than Oja's rule: there is a hyperplane—i.e., a continuum—in synaptic weight space of stable weight configurations for which the homeostatic condition of postsynaptic firing rate reaching a target is achieved. This opens the possibility that a more complex model of the circuit could allow consolidation of an analog memory using the BCM rule. However, such a rule would not allow for the weight of a single input type to change independently of the others, unlike the heterosynaptic rule.

S3 Stability of feedforward model with closed-loop feedback about errors

In the main text (Materials and methods, *Analysis of synaptic learning dynamics*), we saw that, for the feedforward model, during the post-training period all trajectories tended toward a fixed point along a line attractor. Here we will show that during training, closed-loop feedback about behavioral errors, carried by the climbing fibers, can stably drive the circuit to a fixed point that fully reduces oculomotor errors on average (Fig. S1).

We assume that the dynamics of w_H and v are given by the learning rules Eqs. (19) and (20). During training, we assume that the climbing fiber modulation $\delta\mathrm{CF}(t)$ is modeled by Eq. (12), i.e., it is a saturating function of the oculomotor (retinal slip) error $\dot{R}(t)$ (Eq. (11)). Then, there is a fixed point of the weight dynamics corresponding to the point at which the measured gain of the

system is $g = g^{\text{target}}$,

$$(w_H^*, v^*) = \left(0, \frac{g^{\text{target}}}{k_E k_{\text{MF}}}\right). \tag{S.13}$$

The dynamics of trajectories leading to a steady state can be separated into two regimes. Initially, the amplitude of oculomotor errors is large, so that δCF saturates, and correspondingly so does the parallel fiber-climbing fiber covariance $c(t)$. As a result, the weight w_H tends toward the value defined by Eq. (21). From the simplified form of the heterosynaptic plasticity rule for v, Eq. (20), this leads to initially linear growth in v. As v increases, eventually the amplitude of oculomotor errors will decrease, such that δCF will be a linear function of $\dot{R}$, and from Eq. (23), $c(t)$ will be a linear function of w_H and v, such that

$$\tau_w \frac{\mathrm{d}w_H}{\mathrm{d}t} \approx -w_H + k_{\text{LTD}} k_{\text{PF}}^2 k_{\text{CF}} \beta_{\text{light}} k_E w_{\text{PC}} \left(\frac{g^{\text{target}} - k_E k_{\text{MF}} v}{k_E k_{\text{PF}} w_{\text{PC}}} - w_H\right) \sigma_H^2. \tag{S.14}$$

Thus, in the linear regime, the dynamics have eigenvalues

$$\lambda_\pm = -\frac{1}{2\tau_w}(1 + k_{\text{LTD}} k_{\text{PF}}^2 k_{\text{CF}} \beta_{\text{light}} \sigma_H^2 k_E w_{\text{PC}}) \cdot \left[1 \mp \sqrt{1 - \frac{4\tau_w k_{\text{LTD}} k_{\text{PF}} k_{\text{CF}} \beta_{\text{light}} \sigma_H^2 k_E k_{\text{MF}} k_{v,\text{hetero}} \langle \text{MF}(t)\text{PF}(t)\rangle_{\tau_{f,v}}}{(1 + k_{\text{LTD}} k_{\text{PF}}^2 k_{\text{CF}} \beta_{\text{light}} \sigma_H^2 k_E w_{\text{PC}})^2}}\right]. \tag{S.15}$$

Since all of the parameters are strictly positive (during training, $\sigma_H^2 > 0$), the steady state in Eq. (S.13) is a stable fixed point of the dynamics.

The approach to the fixed point could involve decaying oscillations in the values of w_H and v around the steady state. Oscillations do not occur if

$$k_{v,\text{hetero}} \langle \text{MF}(t)\text{PF}(t)\rangle_{\tau_{f,v}} \cdot \tau_w < \frac{(1 + k_{\text{LTD}} k_{\text{PF}}^2 k_{\text{CF}} \beta_{\text{light}} \sigma_H^2 k_E w_{\text{PC}})^2}{4 k_{\text{LTD}} k_{\text{PF}} k_{\text{CF}} \beta_{\text{light}} \sigma_H^2 k_E k_{\text{MF}}}. \tag{S.16}$$

The above inequality places a constraint on the ratio of the rate of plasticity at v relative to the intrinsic rate of plasticity ($1/\tau_w$) at w_H to avoid oscillations.

The above analysis shows that, even when the size of climbing fiber responses decreases across learning as a result of decreased errors in the output, stable consolidation is still achieved. Learning at the early-learning site decreases in amplitude as consolidation occurs at the late-learning site until the system reaches desired performance, at which point the early-learning area is no longer required (Fig. S1).

S4 Analysis of drift in feedforward model without error feedback

In the main text, we noted that, in the feedforward model, consolidation could be understood as temporal integration at the late-learning site v of weight changes initially induced in the early-learning weight w_H. Drawing on other work on integrators, we showed that periodically applied

random perturbations of w_H led to the value of v random-walking over time (Fig. 3*A–D*). Here, we derive analytically a formula for the variance across sample paths of the late-learning weight v as a function of time and the learning rate at v.

We start from the analysis in the Materials and methods (*Analysis of synaptic learning dynamics*). As in the simulation of Figure 3*A,B* (see *Simulation of oculomotor learning*), we assumed that the value of w_H was perturbed at a regular interval T, with each perturbation k given by an i.i.d. random value $\Delta w_H^{(k)}$ drawn uniformly over the interval $[-\alpha, \alpha]$. After the perturbation, the synaptic dynamics cause w_H to decay toward a baseline value of zero. To model the effect of each perturbation, we assume the value of w_H immediately following the kth perturbation is given by

$$w_{H,\text{perturb}}^{(k)} = w_{H,\text{post}}^{(k-1)} + \Delta w_H^{(k)}, \tag{S.17}$$

where the value of w_H just prior to the kth perturbation is given by

$$w_{H,\text{post}}^{(k-1)} = w_{H,\text{perturb}}^{(k-1)} e^{-T/\tau_w}, \tag{S.18}$$

which models the decay in w_H between perturbations. We can simplify this recurrence relation as

$$\begin{aligned} w_{H,\text{post}}^{(k)} &= e^{-T/\tau_w}(w_{H,\text{post}}^{(k-1)} + \Delta w_H^{(k)}) \\ &= e^{-T/\tau_w}(e^{-T/\tau_w}(w_{H,\text{post}}^{(k-2)} + \Delta w_H^{(k-1)}) + \Delta w_H^{(k)}) \\ &= w_{\text{post}}^{(-1)}(e^{-T/\tau_w})^{k+1} + \sum_{k'=0}^{k} \Delta w_H^{(k-k')}(e^{-T/\tau_w})^{k'+1} \\ &= w_{\text{post}}^{(-1)}(e^{-T/\tau_w})^{k+1} + \sum_{k'=0}^{k} \Delta w_H^{(k')}(e^{-T/\tau_w})^{k-k'+1}. \end{aligned} \tag{S.19}$$

We assume that the system starts in a stable steady state, i.e., $w_{H,\text{post}}^{(-1)} = 0$. The change during each post-perturbation period $k \geq 0$ is thus

$$\begin{aligned} \Delta w_{H,\text{post}}^{(k)} &= w_{H,\text{post}}^{(k)} - w_{H,\text{perturb}}^{(k)} = w_{H,\text{perturb}}^{(k)}(e^{-T/\tau_w} - 1) \\ &= \left(\sum_{k'=0}^{k-1} \Delta w_H^{(k')}(e^{-T/\tau_w})^{k-k'} + \Delta w_H^{(k)}\right)(e^{-T/\tau_w} - 1) \\ &= (e^{-T/\tau_w} - 1)\sum_{k'=0}^{k} \Delta w_H^{(k')}(e^{-T/\tau_w})^{k-k'}. \end{aligned} \tag{S.20}$$

Then, from Eq. (25), the change in v during the kth perturbation period can be approximated as

$$\begin{aligned} \Delta v^{(k)} &\approx k_{v,\text{hetero}}\tau_w \langle \text{MF}(t)\text{PF}(t)\rangle_{\tau_{f,v}} \Delta w_{H,\text{post}}^{(k)} \\ &= -k_{v,\text{hetero}}\tau_w \langle \text{MF}(t)\text{PF}(t)\rangle_{\tau_{f,v}} (1 - e^{-T/\tau_w}) \sum_{k'=0}^{k} \Delta w_H^{(k')}(e^{-T/\tau_w})^{k-k'}, \end{aligned} \tag{S.21}$$

using the relationship in Eq. (S.20).

We can calculate the overall value of v after the kth perturbation as

$$v^{(k)} = v_0 + \sum_{k'=0}^{k} \Delta v^{(k')}. \tag{S.22}$$

This has variance

$$\begin{aligned}
\mathrm{Var}(v^{(k)}) &= \mathrm{Var}\left(v_0 + \sum_{k'=0}^{k} \Delta v^{(k')}\right) = \mathrm{Var}\left(\sum_{k'=0}^{k} \Delta v^{(k')}\right) \\
&= \sum_{k'=0}^{k} \sum_{k''=0}^{k} \mathrm{Cov}(\Delta v^{(k')}, \Delta v^{(k'')}) \\
&= \sum_{k'=0}^{k} \left(2 \sum_{k''=0}^{k'} \mathrm{Cov}(\Delta v^{(k')}, \Delta v^{(k'')}) - \mathrm{Var}(\Delta v^{(k')})\right).
\end{aligned} \tag{S.23}$$

Thus, to calculate the variance in $v^{(k)}$, we need to calculate the variance of $\Delta v^{(k)}$ and the covariance between $\Delta v^{(k)}$ and $\Delta v^{(j)}$. For convenience, define

$$E = e^{-T/\tau_w}, \text{ and} \tag{S.24}$$

$$K = k_{v,\mathrm{hetero}} \tau_w \langle \mathrm{MF}(t)\mathrm{PF}(t)\rangle_{\tau_{f,v}}. \tag{S.25}$$

We can write the covariance between the change in v due to perturbations j and k as

$$\begin{aligned}
\mathrm{Cov}(\Delta v^{(k)}, \Delta v^{(j)}) &= (K(1-E))^2 \sum_{k'=0}^{\min(k,j)} \mathrm{Var}(\Delta w_H^{(k')}) E^{(k+j-2k')} \\
&= \frac{1}{3}\alpha^2 (K(1-E))^2 \sum_{k'=0}^{\min(k,j)} E^{(k+j-2k')}.
\end{aligned} \tag{S.26}$$

where, because $\Delta w_H^{(k)}$ is uniformly distributed over $[-\alpha, \alpha]$, $\mathrm{Var}(\Delta w_H^{(k)}) = 1/3 \cdot \alpha^2$. The variance of $\Delta v^{(k)}$ can be found by taking $j = k$ in Eq. (S.26),

$$\begin{aligned}
\mathrm{Var}(\Delta v^{(k)}) = \mathrm{Cov}(\Delta v^{(k)}, \Delta v^{(k)}) &= \frac{1}{3}\alpha^2 (K(1-E))^2 \sum_{k'=0}^{k} E^{2k'} \\
&= (K(1-E))^2 \cdot \frac{1}{3}\alpha^2 \cdot \frac{1-E^{2(k+1)}}{1-E^2} \\
&= \frac{1}{3}\alpha^2 K^2 \cdot \frac{1-E}{1+E}(1-E^{2(k+1)}).
\end{aligned} \tag{S.27}$$

We can then calculate the sum

$$\sum_{k'=0}^{k} \mathrm{Var}(\Delta v^{(k')}) = \frac{1}{3}\alpha^2 K^2 \frac{1-E}{1+E} \sum_{k'=0}^{k} (1 - E^{2(k'+1)})$$
$$= \frac{1}{3}\alpha^2 K^2 \left(\frac{1-E}{1+E}(k+1) - \frac{E^2(1-E^{2(k+1)})}{(1+E)^2} \right) \tag{S.28}$$

Finally, from (S.26), we can also calculate the sum of covariances

$$\sum_{k'=0}^{k} \sum_{k''=0}^{k'} \mathrm{Cov}(\Delta v^{(k')}, \Delta v^{(k'')}) = \frac{1}{3}\alpha^2 (K(1-E))^2 \sum_{k'=0}^{k} \sum_{k''=0}^{k'} \sum_{k'''=0}^{k''} E^{(k'+k''-2k''')}$$
$$= \frac{1}{3}\alpha^2 K^2 \left(\frac{k+1}{1+E} - \frac{E(1-E^{k+1})}{1-E} + \frac{E^3(1-E^{2(k+1)})}{(1-E)(1+E)^2} \right). \tag{S.29}$$

Putting these results together, from Eq. (S.23), the variance of $v^{(k)}$ is

$$\mathrm{Var}(v^{(k)}) = \sum_{k'=0}^{k} \left(2 \sum_{k''=0}^{k'} \mathrm{Cov}(\Delta v^{(k')}, \Delta v^{(k'')}) - \mathrm{Var}(\Delta v^{(k')}) \right)$$
$$= \frac{1}{3}\alpha^2 K^2 \Bigg(-\frac{1-E}{1+E}(k+1) + \frac{E^2(1-E^{2(k+1)})}{(1+E)^2}$$
$$+ 2 \left[\frac{k+1}{1+E} - \frac{E(1-E^{k+1})}{1-E} + \frac{E^3(1-E^{2(k+1)})}{(1-E)(1+E)^2} \right] \Bigg)$$
$$= \frac{1}{3}\alpha^2 K^2 \left(k + 1 - \frac{2E(1-E^{k+1})}{1-E} + \frac{E^2(1-E^{2(k+1)})}{(1-E^2)} \right). \tag{S.30}$$

That is, the term $A(k)$ in Eq. (28) in the main methods is given by

$$A(k) = -\frac{2E(1-E^{k+1})}{1-E} + \frac{E^2(1-E^{2(k+1)})}{(1-E^2)}. \tag{S.31}$$

S5 Analysis of model dynamics during training with fixed-strength internal feedback loop

In the main text, we proposed two circuit mechanisms that allow for consolidation to occur successfully when the circuit contains an internal feedback loop from the late- to the early-learning area. Here we analyze the dynamics of the synaptic weights in the circuit during training for each mechanism, and show that under similar conditions as post-training, both mechanisms are also stable during training and result in the circuit correctly learning the target input-to-output computation (Fig. S4).

S5.1 Local stability of the fixed point during training

We first show that both mechanisms have the same locally stable fixed point, corresponding to the circuit achieving the target gain. For both mechanisms, the learning rule for the late-learning weight v is identical, Eq. (58), and the only difference comes from the effect on the early-learning feedforward weight w_H. We assume that the weight of the feedback pathway is fixed and that the total feedback strength is less than 1, as in Eq. (62). We also assume that $w^+_{H,\infty} = 0$ and that during training the variability of the head input is $\langle \dot{H}^2 \rangle_{\tau_f} = \sigma_H^2 > 0$. In our analysis, we make use of the fact that, from Eq. (52), the error in gain can be expanded as

$$g^{\text{target}} - g(t) = -\frac{k_E w_{\text{PC}} k_{\text{PF},H}}{1 - w_{\text{PC}} k_E \tilde{w}_E} (w^*_{H,\text{opt}}(v(t)) - w_H(t)), \tag{S.32}$$

and for convenience we use variables with tildes as defined in Eqs. (53)–(55), and where

$$w^*_{H,\text{opt}}(v) = -\frac{1}{w_{\text{PC}} k_{\text{PF},H}} \left(\frac{g^{\text{target}}}{k_E} (1 - w_{\text{PC}} k_E \tilde{w}_E) - k_{\text{MF}} v \right) \tag{S.33}$$

is the value of w_H that would minimize the error in gain for a fixed choice of v. Similarly, we define the value of w_H that would reset early-learning activity for a fixed choice of v as

$$w^*_{H,\text{reset}}(v) = \frac{k_E \tilde{w}_E k_{\text{MF}} v}{k_{\text{PF},H}}. \tag{S.34}$$

For the inhibitory plasticity mechanism (Fig. 6*A*), changes in the excitatory weight w_H^+ are governed by instructive climbing fiber input driven by the retinal slip signal, Eq. (12), according to the learning rule in Eq. (56), whereas the inhibitory weight w_H^- tends to drive the activity of the early-learning area back to its long-term average, as in Eq. (60). More precisely, for a target gain value g^{target}, plasticity at w_H^+ is driven by the parallel fiber-climbing fiber covariance

$$\begin{aligned} c_H(t) &= k_{\text{LTD}} \langle \delta \text{PF}_H(t) \delta \text{CF}(t) \rangle_{\tau_{f,w}} \\ &= k_{\text{LTD}} k_{\text{CF}} k_{\text{PF},H} \langle \dot{H}(t) \tanh[\beta_{\text{light}} (g^{\text{target}} - g(t)) \dot{H}(t)] \rangle_{\tau_{f,w}}. \end{aligned} \tag{S.35}$$

We can combine the learning rules for w_H^+ and the inhibitory weight w_H^- to write an expression for the change in the net weight $w_H = w_H^+ - w_H^-$, and simplify using the same logic as for the post-training period (see Materials and methods, *Analysis of synaptic learning dynamics*). This yields

$$\begin{aligned} \frac{\mathrm{d}w_H}{\mathrm{d}t} \approx k_{\text{LTD}} k_{\text{CF}} k_{\text{PF},H} \left\langle \dot{H} \tanh \left[\beta_{\text{light}} \frac{k_E w_{\text{PC}} k_{\text{PF},H}}{1 - w_{\text{PC}} k_E \tilde{w}_E} (w^*_{H,\text{opt}}(v) - w_H) \dot{H} \right] \right\rangle_{\tau_{f,w}} \\ + \frac{k_{\text{inh}} k_{\text{PF},H}^2 \sigma_H^2}{1 - w_{\text{PC}} k_E \tilde{w}_E} (w^*_{H,\text{reset}}(v) - w_H), \end{aligned} \tag{S.36}$$

using the simplifications in Eqs. (S.32)–(S.34).

For the inhibition of instructive signals mechanism (Fig. 6*F*), the climbing fiber input is both driven by retinal slip error and inhibited by a pathway carrying Purkinje cell output, Eq. (68), leading to a parallel fiber-climbing fiber covariance

$$c_H(t) = k_{\mathrm{LTD}} k_{\mathrm{CF}} k_{\mathrm{PF},H} \langle \dot{H}(t) \tanh[\beta_{\mathrm{light}}(g^{\mathrm{target}} - g(t))\dot{H}(t) + \beta_{\mathrm{reset}}\delta\mathrm{PC}(t)]\rangle_{\tau_{f,w}}. \tag{S.37}$$

This yields a simplified learning rule for w_H:

$$\frac{\mathrm{d}w_H}{\mathrm{d}t} \approx k_{\mathrm{LTD}} k_{\mathrm{CF}} k_{\mathrm{PF},H} \left\langle \dot{H} \tanh\left[\left(\beta_{\mathrm{light}} \frac{k_E w_{\mathrm{PC}} k_{\mathrm{PF},H}}{1 - w_{\mathrm{PC}} k_E \tilde{w}_E}(w^*_{H,\mathrm{opt}}(v) - w_H) + \frac{\beta_{\mathrm{reset}} k_{\mathrm{PF},H}}{1 - w_{\mathrm{PC}} k_E \tilde{w}_E}(w^*_{H,\mathrm{reset}}(v) - w_H)\right)\dot{H}\right]\right\rangle_{\tau_{f,w}}. \tag{S.38}$$

Both mechanisms have the same fixed point,

$$(w^*_H, v^*) = \left(\frac{k_{\mathrm{PF},E} w_E g^{\mathrm{target}}}{k_{\mathrm{PF},H}}, \frac{g^{\mathrm{target}}}{k_E k_{\mathrm{MF}}}\right), \tag{S.39}$$

at which point the gain is equal to g^{target} and $w^*_H = w^*_{H,\mathrm{opt}}(v^*) = w^*_{H,\mathrm{reset}}(v^*)$. Furthermore, the fixed point lies along the post-training line attractor, Eq. (59). Close to the fixed point, the dynamics for both mechanisms are linear, as the hyperbolic tangent term is in an approximately linear regime,

$$\frac{\mathrm{d}w_H}{\mathrm{d}t} \approx -\frac{1}{\tau_{w_H,\mathrm{light}}(\sigma_H^2)}[w_H - (\alpha w^*_{H,\mathrm{opt}}(v) + (1-\alpha) w^*_{H,\mathrm{reset}}(v))]. \tag{S.40}$$

where we defined

$$\alpha = \frac{k_{\mathrm{light}}}{k_{\mathrm{light}} + k_{\mathrm{reset}}}, \text{ and} \tag{S.41}$$

$$\tau_{w_H,\mathrm{light}}(\sigma_H^2) = \frac{1 - w_{\mathrm{PC}} k_E \tilde{w}_E}{(k_{\mathrm{light}} + k_{\mathrm{reset}}) k_{\mathrm{PF},H}^2 \sigma_H^2} = (1-\alpha)\tau_{w_H,\mathrm{dark}}(\sigma_H^2), \tag{S.42}$$

with

$$k_{\mathrm{light}} = k_{\mathrm{LTD}} k_{\mathrm{CF}} \beta_{\mathrm{light}} k_E w_{\mathrm{PC}}, \tag{S.43}$$

and

$$k_{\mathrm{reset}} = \begin{cases} k_{\mathrm{inh}} & \text{for the inhibitory plasticity mechanism} \\ k_{\mathrm{LTD}} k_{\mathrm{CF}} \beta_{\mathrm{reset}} & \text{inhibition of instructive signals} \end{cases}. \tag{S.44}$$

From Eqs. (58) and (S.40), the Jacobian is

$$\mathbf{J} = -\frac{\sigma_H^2}{1 - w_{\mathrm{PC}} k_E \tilde{w}_E} \cdot \begin{pmatrix} (k_{\mathrm{light}} + k_{\mathrm{reset}}) k_{\mathrm{PF},H}^2 & -k_{\mathrm{MF}} k_{\mathrm{PF},H} k_E (k_{\mathrm{reset}} \tilde{w}_E + k_{\mathrm{LTD}} k_{\mathrm{CF}} \beta_{\mathrm{light}}) \\ k_{v,\mathrm{hetero}} k_{\mathrm{MF}} k_{\mathrm{PF},H} & -k_{v,\mathrm{hetero}} k_{\mathrm{MF}}^2 k_E \tilde{w}_E \end{pmatrix}. \tag{S.45}$$

The eigenvalues of the Jacobian are

$$\lambda_{\pm} = \frac{1}{2}\left(\frac{1}{\tau_v} - \frac{1}{\tau_{w_H,\text{light}}} \pm \left|\frac{1}{\tau_v} - \frac{1}{\tau_{w_H,\text{light}}}\right|\sqrt{1 - \frac{4k_{v,\text{hetero}}k_{\text{MF}}^2 k_{\text{PF},H}^2 k_E k_{\text{LTD}} k_{\text{CF}} \beta_{\text{light}} \sigma_H^4}{(1 - w_{\text{PC}} k_E \tilde{w}_E)(1/\tau_v - 1/\tau_{w_H,\text{light}})^2}}\right), \quad \text{(S.46)}$$

where for conciseness, we omitted that $\tau_{w_H,\text{light}}$ and τ_v are functions of σ_H^2, with the latter defined as in Eq. (66). Since by definition all parameter values are positive, during training $\sigma_H^2 > 0$, and we assumed that Eq. (62) holds, both eigenvalues have negative real part and the fixed point is stable if

$$\frac{k_{v,\text{hetero}} k_{\text{MF}}^2 k_E \tilde{w}_E}{(k_{\text{light}} + k_{\text{reset}}) k_{\text{PF},H}^2} = \frac{\tau_{w_H,\text{light}}(\sigma_H^2)}{\tau_v(\sigma_H^2)} < 1. \quad \text{(S.47)}$$

Note that if the post-training stability condition, Eq. (64), is met, then this stability condition is also automatically met, because by definition $\tau_{w_H,\text{dark}} > \tau_{w_H,\text{light}}$ for all $\sigma_H^2 > 0$ (see Eq. (S.42)).

Outside of the region near the fixed point, the dynamics of the weights differ between the two mechanisms. In the following sections, we evaluate the stability of the dynamics in this region for each mechanism.

S5.2 Inhibitory plasticity-driven resetting

For the inhibitory plasticity mechanism, we can understand the dynamics by first looking at w_H, as defined by the learning rule in Eq. (S.36). During training to increase the gain, w_H will both be driven by the climbing fiber input away from its initial value along the post-training attractor, towards the line of points $(w_{H,\text{opt}}^*(v), v)$ (along which $g = g^{\text{target}}$), and by the inhibitory plasticity back towards the attractor (the v-nullcline). Thus, the nullcline for w_H lies in between these two lines. As training continues, the weight trajectory will cross the w_H-nullcline, entering a region in which both v and w_H are increasing. If the slope of the flow field along the $w_{H,\text{opt}}^*$ line is shallower than the slope of the line itself, then the trajectory will stay bounded between the $w_{H,\text{opt}}^*$ line and the w_H-nullcline, reaching the fixed point without any overshoot in the circuit gain (Fig. S4A,B). Otherwise, the trajectory will cross the $w_{H,\text{opt}}^*$ line, causing an overshoot in the gain. In this case, the trajectory will still tend toward the fixed point as long as the post-training stability condition, Eq. (64), holds.

First, we determine the conditions under which the trajectory will reach the fixed point without any overshoot in gain. The slope of the flow field along the $w_{H,\text{opt}}^*$ line is

$$\left.\frac{\mathrm{d}v}{\mathrm{d}w_H}\right|_{(w_{H,\text{opt}}^*(v),v)} = \frac{k_{v,\text{hetero}} k_{\text{MF}}}{k_{\text{inh}} k_{\text{PF},H}} = \left.\frac{\mathrm{d}v}{\mathrm{d}w_H}\right|_{\text{post}}, \quad \text{(S.48)}$$

i.e., it is equal to the slope of trajectories during the post-training period (see Eq. (63)). The slope

of the $w^*_{H,\text{opt}}$ line is

$$\frac{\mathrm{d}}{\mathrm{d}w_H}(v^*_{\text{opt}}(w_H)) = \frac{k_{\text{PF},H} w_{\text{PC}}}{k_{\text{MF}}} = w_{\text{PC}} k_E \tilde{w}_E \cdot \frac{k_{\text{PF},H}}{k_E \tilde{w}_E k_{\text{MF}}}, \tag{S.49}$$

where we rearrange Eq. (S.33) to define $v^*_{\text{opt}}(w_H)$ as the value of v that satisfies $w^*_{H,\text{opt}} = w_H$. Note that the second term in the rightmost expression above is the slope of the v-nullcline (i.e., the post-training attractor). Therefore, following the reasoning leading to the post-training stability condition in Eq. (64), the slope of the flow field along the $w^*_{H,\text{opt}}$ line will be shallower than the slope of the line itself if

$$\frac{\tau_{w_H,\text{dark}}(\sigma_H^2)}{\tau_v(\sigma_H^2)} < w_{\text{PC}} k_E \tilde{w}_E, \tag{S.50}$$

which can also be written as

$$\frac{k_{v,\text{hetero}} k_{\text{MF}}^2}{k_{\text{reset}} k_{\text{PF},H}^2} < w_{\text{PC}}. \tag{S.51}$$

Note that, since $w_{\text{PC}} k_E \tilde{w}_E < 1$ (Eq. (62)), this is a stricter condition than Eq. (64).

If the condition in Eq. (S.50) does not hold, then the trajectory will cross the $w^*_{H,\text{opt}}$ line. In this region, from Eqs. (58) and (S.32)–(S.36), the instantaneous slope of the trajectory is given by

$$\begin{aligned}\frac{\mathrm{d}v}{\mathrm{d}w_H} &\approx \frac{k_{v,\text{hetero}} k_{\text{PF},H} k_{\text{MF}} \sigma_H^2 (w^*_{H,\text{reset}} - w_H)}{k_{\text{inh}} k_{\text{PF},H}^2 \sigma_H^2 (w^*_{H,\text{reset}} - w_H) - c_H(t)(1 - w_{\text{PC}} k_E \tilde{w}_E)} \\ &= \frac{1}{1 - \frac{c_H(t)(1 - w_{\text{PC}} k_E \tilde{w}_E)}{k_{\text{inh}} k_{\text{PF},H} \sigma_H^2 (w^*_{H,\text{reset}} - w_H)}} \cdot \left.\frac{\mathrm{d}v}{\mathrm{d}w_H}\right|_{\text{post}},\end{aligned} \tag{S.52}$$

where $c_H(t)$ is as defined in Eq. (S.35). Since c_H has a saturating nonlinearity, $-C k_{\text{LTD}} k_{\text{CF}} k_{\text{PF},H} < c_H \leq 0$ in the region $w_H \leq w^*_{H,\text{opt}} \leq w^*_{H,\text{reset}}$, with C as defined by Eq. (22). Then, the slope of trajectories is bounded,

$$\frac{1}{1 + \frac{C k_{\text{LTD}} k_{\text{CF}} k_{\text{PF},H} (1 - w_{\text{PC}} k_E \tilde{w}_E)}{k_{\text{inh}} k_{\text{PF},H} \sigma_H^2 (w^*_{H,\text{reset}} - w_H)}} \cdot \left.\frac{\mathrm{d}v}{\mathrm{d}w_H}\right|_{\text{post}} < \frac{\mathrm{d}v}{\mathrm{d}w_H} \leq \left.\frac{\mathrm{d}v}{\mathrm{d}w_H}\right|_{\text{post}}, \tag{S.53}$$

where the right-hand equality holds along the $w^*_{H,\text{opt}}$ line, as we saw above. Therefore, as long as Eq. (64) holds (i.e., the slope of post-training trajectories is shallower than the slope of the v-nullcline), trajectories during training will also have slope shallower than the v-nullcline and tend toward the fixed point.

S5.3 Inhibition of instructive signals-driven resetting

For the inhibition of instructive signals mechanism, in the main text (Materials and methods, *Reset Mechanism 2: Inhibition of pathway carrying instructive signals*) we saw that post-training dynamics could become unstable if the change in w_H during the training period was too large,

limiting the maximum change in gain that the circuit could stably consolidate. This was defined by the slope of trajectories in the region of weight space where the climbing fiber response was saturated, Eq. (71), from which we found a stable region of weight space, Eq. (72). Assuming that before training the weights start at a fixed point and that during training the learned change in gain initially only comes from changes in w_H, the largest gain change that could be learned is

$$\begin{aligned}\Delta g_{\max} &= -\frac{w_{\mathrm{PC}} k_E k_{\mathrm{PF},H}}{1 - w_{\mathrm{PC}} k_E \tilde{w}_E} \Delta w_{H,\max} \\ &= -\frac{w_{\mathrm{PC}} k_{\mathrm{PF},H}^2 k_{\mathrm{LTD}} k_{\mathrm{CF}} C}{k_{v,\mathrm{hetero}} k_{\mathrm{MF}}^2 \sigma_H^2 \tilde{w}_E},\end{aligned} \tag{S.54}$$

where $\Delta w_{H,\max}$ has value equal to the right hand side of Eq. (72). With this largest gain change in mind, below we examine the weight dynamics during training.

From Eq. (S.38) we can see that for a fixed value of v, $\mathrm{d}w_H/\mathrm{d}t = 0$ when

$$\begin{aligned}w_H &= \frac{\beta_{\mathrm{light}} k_E w_{\mathrm{PC}}}{\beta_{\mathrm{light}} k_E w_{\mathrm{PC}} + \beta_{\mathrm{reset}}} w^*_{H,\mathrm{opt}}(v) + \frac{\beta_{\mathrm{reset}}}{\beta_{\mathrm{light}} k_E w_{\mathrm{PC}} + \beta_{\mathrm{reset}}} w^*_{H,\mathrm{reset}}(v) \\ &= \alpha w^*_{H,\mathrm{opt}}(v) + (1-\alpha) w^*_{H,\mathrm{reset}}(v),\end{aligned} \tag{S.55}$$

where α is defined as in Eq. (S.41). That is, the w_H-nullcline is linear and lies between the line $(w^*_{H,\mathrm{opt}}(v), v)$ (along which $g = g^{\mathrm{target}}$) and the post-training attractor (v-nullcline). Thus, during training to increase the gain, initially w_H will decrease from a point on the post-training attractor toward the w_H-nullcline, during which time v will start to grow. The trajectory will then cross the w_H-nullcline. As long as the trajectory does so within the region defined by Eq. (72), the trajectory will tend toward the post-training attractor and eventually the fixed point (Fig. S4*D*).

However, it is possible that trajectories may cross the $w^*_{H,\mathrm{opt}}$ line, which would correspond to the gain of the circuit overshooting the target. We can try to ensure that trajectories do not overshoot by keeping them within the region bounded by $w^*_{H,\mathrm{opt}}$ and the w_H-nullcline. If the $w^*_{H,\mathrm{opt}}$ line is near the w_H-nullcline and therefore within the linear regime of the climbing fiber response δCF, the slope of trajectories along the line is approximately as defined in Eq. (S.49). As for the inhibitory plasticity reset mechanism, trajectories will therefore stay approximately within the region between the $w^*_{H,\mathrm{opt}}$ line and the w_H-nullcline if the condition in Eq. (S.50) is met. If the $w^*_{H,\mathrm{opt}}$ line is far from the nullcline, and the climbing fiber response is saturated, then the slope of the flow field along the $w^*_{H,\mathrm{opt}}$ line is

$$\left.\frac{\mathrm{d}v}{\mathrm{d}w_H}\right|_{(w^*_{H,\mathrm{opt}}(v),v)} \approx \frac{k_v k_{\mathrm{MF}} \sigma_H^2}{k_{\mathrm{LTD}} k_{\mathrm{CF}} C (1 - w_{\mathrm{PC}} k_E \tilde{w}_E)} (w^*_{H,\mathrm{reset}} - w^*_{H,\mathrm{opt}}). \tag{S.56}$$

Then, trajectories will stay bounded by the $w^*_{H,\text{opt}}$ line and there will be no overshoot in gain if

$$\left.\frac{\mathrm{d}v}{\mathrm{d}w_H}\right|_{(w^*_{H,\text{opt}}(v),v)} < \frac{\mathrm{d}}{\mathrm{d}w_H}(v^*_{\text{opt}}),$$

where $\mathrm{d}/\mathrm{d}w_H(v^*_{\text{opt}})$ is the slope of the $w^*_{H,\text{opt}}$ line defined in Eq. (S.49). This directly gives that, for a given value of v, the distance between the $w^*_{H,\text{opt}}$ line and the v-nullcline $w^*_{H,\text{reset}}(v)$ can be no larger than

$$w^*_{H,\text{reset}}(v) - w^*_{H,\text{opt}}(v) < w_{\text{PC}} k_E \tilde{w}_E \Delta w_{H,\max}. \tag{S.57}$$

S6 Circuit model with plastic internal feedback

In this section, we extend the feedback model discussed in the previous section and in the main text (**Materials and methods**, *Circuit model with internal feedback loop*) to the case in which the weight of the internal feedback pathway to the early-learning area is also plastic. Plasticity in the excitatory feedforward and feedback weights, w^+_H and w^+_E respectively, is described by Eq. (56). The plasticity at the late-learning site v is described, as before, by Eq. (58). Similar to the circuits with fixed strength feedback, we can show how changes at the early-learning site can be successfully consolidated post-training if the circuit either has plasticity at w^-_H (and/or w^-_E), or through inhibition of the climbing fiber instructive signal pathway, both of which reset early-learning activity so that the output of the early-learning area does not modulate in response to sensory input.

Below, we describe in further detail the implementation of each of these reset mechanisms in the circuit. Then, we show that the linearized dynamics of the circuit with either reset mechanism are the same, and determine conditions under which learning and consolidation occur stably. In general, learning of a target input-output gain and post-training consolidation will be locally stable around its fixed point as long as the weight of the feedback pathway w_E stays within the region $w_E < 1/(w_{\text{PC}} k_E k_{\text{PF},E})$, so that the strength of the feedback around the loop from early- to late-learning area and back to the early-learning area is less than 1, and if the rate of plasticity at the early-learning sites is fast compared to the rate of plasticity at the late-learning site.

S6.1 Model formulation

We implemented the two circuit reset mechanisms by extending the model with non-plastic feedback (see Materials and methods, *Circuit model with internal feedback loop*) as follows.

Inhibitory plasticity-driven resetting For the inhibitory plasticity mechanism we model plasticity in w^-_H as governed by Eq. (60) and in w^-_E by

$$\frac{\mathrm{d}w^-_E}{\mathrm{d}t} = k_{\text{inh}} \langle \text{PF}_E(t) \delta\text{PC}(t) \rangle_{\tau_{\text{inh}}}, \tag{S.58}$$

where $\delta\mathrm{PC}(t) = \mathrm{PC}(t) - \langle \mathrm{PC}(t)\rangle_{\tau_f}$. During training with a target gain of g^{target}, plasticity at the excitatory weights to the early-learning area is driven by a climbing fiber signal carrying only retinal slip, Eq. (12). Plasticity at the feedforward excitatory weight w_H^+ is driven by (negative of) the feedforward parallel fiber-climbing fiber covariance $c_H(t)$, Eq. (S.35), leading to the simplified learning rule for the net weight w_H in Eq. (S.36), where now $w^*_{H,\mathrm{opt}}$ and $w^*_{H,\mathrm{reset}}$ are functions of both w_E and v. Similarly, plasticity at w_E^+ is driven by (negative of) the feedback covariance

$$\begin{aligned} c_E(t) &= k_{\mathrm{LTD}}\langle \delta\mathrm{PF}_E(t)\mathrm{CF}(t)\rangle_{\tau_{f,w}} \\ &\approx -k_{\mathrm{LTD}}k_{\mathrm{CF}}k_{\mathrm{PF},E}\cdot g(t)\langle \dot{H}(t)\tanh[\beta_{\mathrm{light}}(g^{\mathrm{target}} - g(t))\dot{H}(t)]\rangle_{\tau_{f,w}}. \end{aligned} \tag{S.59}$$

As for the feedforward input, we note that

$$g^{\mathrm{target}} - g(t) = \frac{w_{\mathrm{PC}}k_E k_{\mathrm{PF},E}\cdot g^{\mathrm{target}}}{1 - w_{\mathrm{PC}}k_E\tilde{w}_E(t)}(w^*_{E,\mathrm{opt}}(w_H(t), v(t)) - w_E(t)), \tag{S.60}$$

where we define

$$w^*_{E,\mathrm{opt}}(w_H, v) = \frac{1}{w_{\mathrm{PC}}k_E k_{\mathrm{PF},E}}\left[1 - \frac{k_E}{g^{\mathrm{target}}}(k_{\mathrm{MF}}v - w_{\mathrm{PC}}k_{\mathrm{PF},H}w_H)\right], \tag{S.61}$$

i.e., the value of w_E that would minimize errors for a fixed choice of w_H and v. Further defining

$$w^*_{E,\mathrm{reset}}(w_H, v) = \frac{k_{\mathrm{PF},H}w_H}{k_E k_{\mathrm{PF},E}k_{\mathrm{MF}}v}, \tag{S.62}$$

we can write a combined learning rule for the net feedback weight w_E,

$$\begin{aligned} \frac{\mathrm{d}w_E}{\mathrm{d}t} \approx \delta_{\mathrm{train}}k_{\mathrm{LTD}}k_{\mathrm{CF}}k_{\mathrm{PF},E}\cdot g\left\langle \dot{H}\tanh\left[\beta_{\mathrm{light}}\frac{w_{\mathrm{PC}}k_E k_{\mathrm{PF},E}\cdot g^{\mathrm{target}}}{1 - w_{\mathrm{PC}}k_E k_{\mathrm{PF},E}w_E}\left(w^*_{E,\mathrm{opt}}(w_H, v) - w_E\right)\dot{H}\right]\right\rangle_{\tau_{f,w}} \\ + \frac{k_{\mathrm{inh}}k^2_{\mathrm{PF},E}\sigma^2_H\cdot g\cdot k_E k_{\mathrm{MF}}v}{1 - w_{\mathrm{PC}}k_E k_{\mathrm{PF},E}w_E}\left(w^*_{E,\mathrm{reset}}(w_H, v) - w_E\right), \end{aligned} \tag{S.63}$$

where $\delta_{\mathrm{train}} = 1$ during training. Post-training, $\delta_{\mathrm{train}} = 0$ since $\dot{R} \equiv 0$, which implies that $c_H(t) = c_E(t) \equiv 0$. Similarly, post-training plasticity at w_H is governed by Eq. (61).

Inhibition of instructive signals-driven resetting For the inhibition of instructive signals mechanism, we assume the inhibitory weight is fixed and that plasticity at both w_H^+ and w_E^+ is governed by the climbing fiber response defined in Eq. (68) (Fig. S3*A*). More specifically, plasticity at w_H^+ is driven by $c_H(t)$ as defined in Eq. (S.37), and we can similarly write the feedback parallel fiber-climbing fiber covariance that drives plasticity at w_E^+ as

$$c_E(t) \approx -k_{\mathrm{LTD}}k_{\mathrm{CF}}k_{\mathrm{PF},E}\cdot g(t)\langle \dot{H}(t)\tanh[\delta_{\mathrm{train}}\beta_{\mathrm{light}}(g^{\mathrm{target}} - g(t))\dot{H}(t) + \beta_{\mathrm{reset}}\delta\mathrm{PC}(t)]\rangle_{\tau_{f,w}}, \tag{S.64}$$

where g^{target} is the target gain of the circuit during training, and $\delta_{\text{train}} = 1$ during training and 0 post-training. From this, we can write that plasticity of the net weight w_E is governed by

$$\frac{\mathrm{d}w_E}{\mathrm{d}t} \approx k_{\text{LTD}} k_{\text{CF}} k_{\text{PF},E} \cdot g \left\langle \dot{H} \tanh\left[\left(\delta_{\text{train}} \beta_{\text{light}} \frac{w_{\text{PC}} k_E k_{\text{PF},E} \cdot g^{\text{target}}}{1 - w_{\text{PC}} k_E k_{\text{PF},E} w_E} \left(w^*_{E,\text{opt}}(w_H, v) - w_E\right) + \beta_{\text{reset}} \frac{k_{\text{PF},E} k_E k_{\text{MF}} v}{1 - w_{\text{PC}} k_E k_{\text{PF},E} w_E} \left(w^*_{E,\text{reset}}(w_H, v) - w_E\right)\right) \dot{H}\right]\right\rangle_{\tau_{f,w}}, \quad \text{(S.65)}$$

where $w^*_{E,\text{opt}}$ and $w^*_{E,\text{reset}}$ are defined by Eqs. (S.61) and (S.62) above.

In Figure S3, we simulated the time evolution of the weights in the circuit implementing this latter mechanism for a 0.5 h period of training to increase the gain to $g^{\text{target}} = 2$ followed by a 23.5 h post-training period. The model parameters were the same as for the model without plasticity of the feedback pathway (Tables 1–3; see Materials and methods, *Simulation of oculomotor learning*).

S6.2 Stability of shared linearized dynamics

Here we extend the analyses of the model without plastic feedback during the training (S5 above) and post-training (Materials and methods, *Circuit model with internal feedback loop*) periods to the case of plastic feedback.

Training We consider the system of differential equations for the three synaptic weights w_H, w_E and v from Eq. (58) and either Eqs. (S.36) and (S.63) for the inhibitory plasticity reset mechanism, or Eqs. (S.38) and (S.65) for the inhibition of instructive signals reset mechanism. For either mechanism, there is a line of steady states (parameterized by s),

$$\begin{pmatrix} w^*_H \\ w^*_E \\ v^* \end{pmatrix} = \begin{pmatrix} k_{\text{PF},E}/k_{\text{PF},H} \cdot g^{\text{target}} \\ 1 \\ 0 \end{pmatrix} s + \begin{pmatrix} 0 \\ 0 \\ g^{\text{target}}/(k_E k_{\text{MF}}) \end{pmatrix}, \quad \text{(S.66)}$$

along which $w^*_H = w^*_{H,\text{opt}} = w^*_{H,\text{reset}}$ and $w^*_E = w^*_{E,\text{opt}} = w^*_{E,\text{reset}}$. This corresponds to the intersection of the surface in weight space that leads the circuit to produce the target gain g^{target} with the surface defined by Eq. (59).

To analyze the stability of the fixed points defined by this line, we calculate the eigenvalues of the Jacobian of the system of differential equations. We note that the Jacobian has the same form for both reset mechanisms after substituting the relevant expressions for k_{reset}, as defined in Eq. (S.44), into the differential equations. Leaving (S.66) parameterized by w^*_E (i.e., taking an arbitrary value of s), we calculate the characteristic polynomial of the Jacobian, which has three roots. One of the roots is $\lambda_1 = 0$ and corresponds to the direction along the line of steady states. Dividing this root out, we are left with $\lambda^2 + a\lambda + b = 0$ for determining the remaining eigenvalues λ_2 and λ_3, where

$$a = -(\lambda_2 + \lambda_3) = \frac{\sigma_H^2}{1 - w_{\mathrm{PC}} k_E k_{\mathrm{PF},E} w_E^*} \cdot \left[(k_{\mathrm{reset}} + k_{\mathrm{light}})(k_{\mathrm{PF},H}^2 + (g^{\mathrm{target}})^2 k_{\mathrm{PF},E}^2) - k_{v,\mathrm{hetero}} k_{\mathrm{MF}}^2 k_E k_{\mathrm{PF},E} w_E^*\right] \quad \text{(S.67)}$$

$$b = \lambda_2 \lambda_3 = k_{v,\mathrm{hetero}} k_E k_{\mathrm{MF}}^2 k_{\mathrm{LTD}} k_{\mathrm{CF}} \beta_{\mathrm{light}} \sigma_H^4 \frac{(g^{\mathrm{target}})^2 k_{\mathrm{PF},E}^2 + k_{\mathrm{PF},H}^2}{1 - w_{\mathrm{PC}} k_E k_{\mathrm{PF},E} w_E^*}, \quad \text{(S.68)}$$

using k_{light} and k_{reset} as defined in Eqs. (S.43) and (S.44). For the eigenvalues to have negative real parts, we must have $a > 0$ and $b > 0$. For $w_E < 1/(w_{\mathrm{PC}} k_E k_{\mathrm{PF},E})$, since all parameters are positive and $\sigma_H^2 > 0$ during training, $b > 0$ is satisfied. $a > 0$ is satisfied if

$$\frac{1}{\tau_{w_H,\mathrm{light}}} + \frac{1}{\tau_{w_E,\mathrm{light}}} > \frac{1}{\tau_v}, \quad \text{(S.69)}$$

where $\tau_{w_H,\mathrm{light}}$ and τ_v are defined as in Eqs. (S.42) and (66) and depend on the value of w_E^* that parameterizes the fixed point as well as on σ_H^2, and where

$$\tau_{w_E,\mathrm{light}}(\sigma_H^2, w_E^*) = \frac{1 - w_{\mathrm{PC}} k_E k_{\mathrm{PF},E} w_E^*}{(k_{\mathrm{reset}} + k_{\mathrm{light}})(g^{\mathrm{target}} k_{\mathrm{PF},E})^2 \sigma_H^2}. \quad \text{(S.70)}$$

That is, for stable learning, the sum of the instantaneous effective rates of plasticity of the early-learning sites w_H and w_E should be faster than the instantaneous effective rate of plasticity at the late-learning site v, extending the result from the fixed-strength feedback case above.

Assuming that w_E^* is upper-bounded by $1/(w_{\mathrm{PC}} k_E k_{\mathrm{PF},E})$, and assuming that the circuit is learning only positive gains so that $g^{\mathrm{target}} > 0$, we can satisfy Eq. (S.69) for the whole range of w_E^* and potential gains with a stricter condition

$$\frac{k_{v,\mathrm{hetero}} k_{\mathrm{MF}}^2}{k_{\mathrm{PF},H}^2 (k_{\mathrm{light}} + k_{\mathrm{reset}})} < w_{\mathrm{PC}}. \quad \text{(S.71)}$$

Post-training In the absence of information about errors, we take $\delta_{\mathrm{train}} = 0$ in Eqs. (S.63) and (S.65). Then, the steady states of the dynamics lie along the surface represented by Eq. (59). Calculating the eigenvalues of the Jacobian at a fixed point $(k_E \tilde{w}_E^* \tilde{v}^* / k_{\mathrm{PF},H}, w_E^*, v^*)$, parameterized by w_E^* and v^*, we find that there are two zero eigenvalues, corresponding to the surface in Eq. (59), and a third eigenvalue

$$\lambda = -\frac{\sigma_H^2}{1 - w_{\mathrm{PC}} k_E \tilde{w}_E^*} \left(k_{\mathrm{reset}}(k_{\mathrm{PF},H}^2 + (k_{\mathrm{PF},E} k_E k_{\mathrm{MF}} v^*)^2) - k_{v,\mathrm{hetero}} k_{\mathrm{MF}}^2 k_E k_{\mathrm{PF},E} w_E^*\right), \quad \text{(S.72)}$$

which is negative if $\sigma_H^2 > 0$, $w_E^* < 1/(w_{\mathrm{PC}} k_E k_{\mathrm{PF},E})$ and

$$\frac{1}{\tau_{w_H,\mathrm{dark}}} + \frac{1}{\tau_{w_E,\mathrm{dark}}} > \frac{1}{\tau_v}, \quad \text{(S.73)}$$

where $\tau_{w_H,\text{dark}}$ and τ_v are defined as in Eqs. (65) and (66) and depend on the value of w_E^* that parameterizes the fixed point as well as on σ_H^2, and where

$$\tau_{w_E,\text{dark}}(\sigma_H^2, w_E^*, v^*) = \frac{1 - w_{\text{PC}} k_E k_{\text{PF},E} w_E^*}{k_{\text{reset}}(k_{\text{PF},E} k_E k_{\text{MF}} v^*)^2 \sigma_H^2}. \tag{S.74}$$

For $w_E^* < 1/(w_{\text{PC}} k_E k_{\text{PF},E})$, and assuming $v^* > 0$, we can write a single stricter condition that is equivalent to Eq. (S.51) to satisfy Eq. (S.73). If this condition is met, then the surface of fixed points (59) will be locally attractive during the post-training period, and since (S.71) will be automatically satisfied, the dynamics during training will be locally stable around the fixed point as well.

References

1. Zenke, F., Hennequin, G. & Gerstner, W. Synaptic Plasticity in Neural Networks Needs Homeostasis with a Fast Rate Detector. *PLoS Computational Biology* **9,** e1003330. doi:10.1371/journal.pcbi.1003330 (2013).
2. Miller, K. D. & MacKay, D. J. C. The Role of Constraints in Hebbian Learning. *Neural Computation* **6,** 100–126. doi:10.1162/neco.1994.6.1.100 (1994).
3. Zenke, F., Gerstner, W. & Ganguli, S. The Temporal Paradox of Hebbian Learning and Homeostatic Plasticity. *Current Opinion in Neurobiology* **43,** 166–176. doi:10.1016/j.conb.2017.03.015 (2017).
4. Abbott, L. F. & Nelson, S. B. Synaptic Plasticity: Taming the Beast. *Nature Neuroscience* **3 Suppl,** 1178–1183. doi:10.1038/81453 (2000).
5. Von der Malsburg, C. Self-Organization of Orientation Sensitive Cells in the Striate Cortex. *Kybernetik* **14,** 85–100. doi:10.1007/BF00288907 (1973).
6. Oja, E. Simplified Neuron Model as a Principal Component Analyzer. *Journal of Mathematical Biology* **15,** 267–273. doi:10.1007/BF00275687 (1982).
7. Turrigiano, G. G. The Self-Tuning Neuron: Synaptic Scaling of Excitatory Synapses. *Cell* **135,** 422–435. doi:10.1016/j.cell.2008.10.008 (2008).
8. Yger, P. & Gilson, M. Models of Metaplasticity: A Review of Concepts. *Frontiers in Computational Neuroscience* **9,** 138. doi:10.3389/fncom.2015.00138 (2015).
9. Tetzlaff, C., Kolodziejski, C., Timme, M. & Wörgötter, F. Synaptic Scaling in Combination with Many Generic Plasticity Mechanisms Stabilizes Circuit Connectivity. *Frontiers in Computational Neuroscience* **5.** doi:10.3389/fncom.2011.00047 (2011).
10. Bienenstock, E. L., Cooper, L. N. & Munro, P. W. Theory for the Development of Neuron Selectivity: Orientation Specificity and Binocular Interaction in Visual Cortex. *Journal of Neuroscience* **2,** 32–48. doi:10.1523/JNEUROSCI.02-01-00032.1982 (1982).

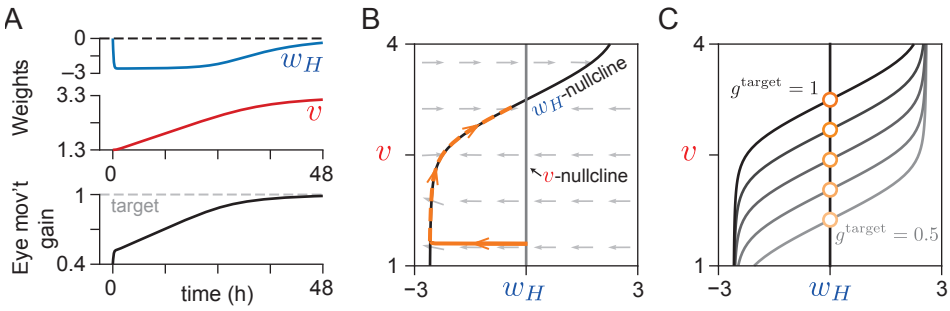

Figure S1: Synaptic weight dynamics of the feedforward-architecture model during training. (A) Dynamics of weights (*top*) and the gain of the circuit (*bottom*) during 48 h of simulated training to increase the sensory input-to-motor output gain of the circuit. (B) Trajectory (orange curve) shows the evolution of synaptic weights in A, which approaches a stable fixed point where the w_H- and v-nullclines (light and dark grey lines) cross. For clarity, the orange trajectory is shown as dashed where it overlaps the w_H-nullcline. (C) The nullclines (grey to black lines) and fixed points (orange circles) of the system for different choices of desired input-to-output gain value (shown by lightness). Any target gain value can be learned by the system, corresponding to a fixed point lying along the line $w_H = 0$. These points are thus also stable when information about behavioral errors is not present.

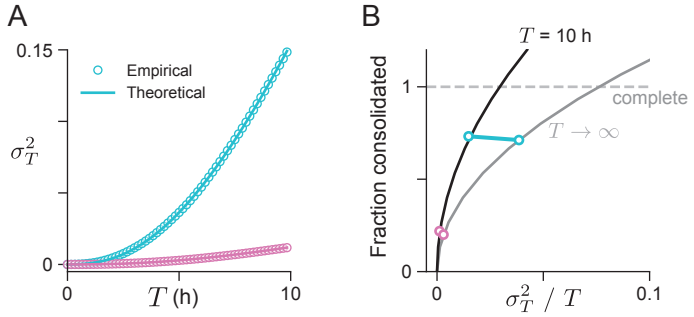

Figure S2: Drift of consolidated memory is asymptotically linear. (A) Comparison of time course of empirical variance of v measured across $N = 250$ simulations (circles; replotted from Fig. 3D) with analytical formula (lines; calculated in S4), for relatively fast (cyan) or slow (magenta) rates of learning at v. (B) Same as Figure 3G, but with tradeoff curve plotted in the asymptotic limit of long time (grey curve).

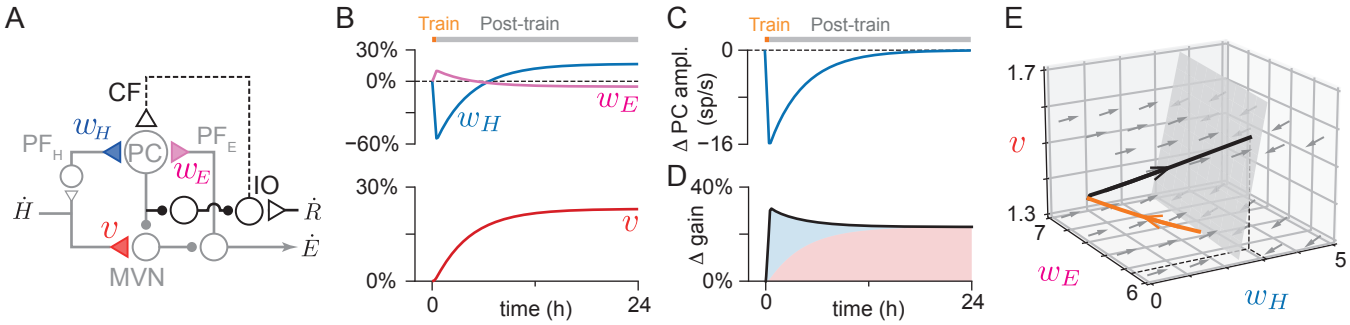

Figure S3: Systems consolidation in a circuit with plastic internal feedback loop. (A) Architecture of circuit with three sites of plasticity, at the feedforward weight onto the early-learning area, w_H (blue), at the weight of the internal feedback pathway, w_E (pink), and at the direct pathway onto the late-learning area, v (red). As in Figure 6F, the post-training reset was achieved by inhibition of the pathway carrying instructive signals to the early-learning area. (B) Change in weights during 0.5 h of simulated training to increase the gain (orange block) followed by 23.5 h post-training with no feedback about behavioral errors (grey block). (C) Early-learning area output (amplitude of Purkinje cell activity relative to moving baseline), which drives consolidation at the late-learning site. (D) Change in the gain of the eye movement response (black line). Blue and red shaded areas show the contributions of the early- and late-learning areas to the circuit transformation. (E) Trajectory of synaptic weights during the training (orange) and post-training (black) periods. Dashed black lines show the projection of the steady state reached at the end of training onto the w_H-w_E plane. Grey arrows show the approximate instantaneous direction in which a weight configuration at a given point in synaptic weight space will evolve during the post-training period, determined analytically. Trajectories tend toward a 2-D surface of marginally stable points (solid grey surface).

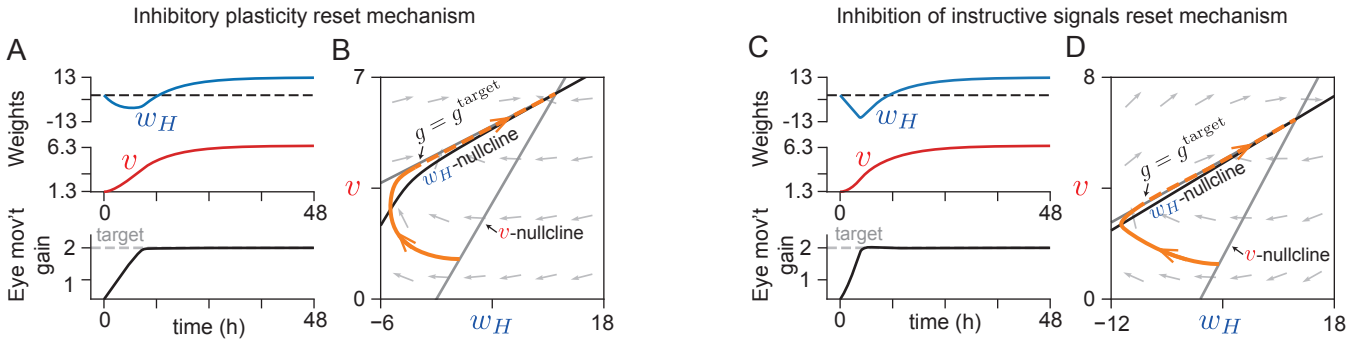

Figure S4: Synaptic weight dynamics during training of the model with feedback architecture. (A) Dynamics of weights (*top*) and the gain of the circuit (*bottom*) during 48 h of simulated training to increase the sensory input-to-motor output gain of the circuit, using the inhibitory plasticity reset mechanism. (B) Trajectory (orange curve) shows the evolution of synaptic weights in A. All trajectories approach a stable fixed point where the w_H- and v-nullclines cross. The second grey line that the trajectory follows corresponds to the constant gain line for the target gain. (C,D) Same as A,B, but for the inhibition of instructive signals reset mechanism. In panels B and D, for clarity, the orange trajectory is shown as dashed where it overlaps the w_H-nullcline.