bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Cerebellar-driven cortical dynamics enable task acquisition, switching and consolidation

Joseph Pemberton[1*], **Paul Chadderton**[2], **Rui Ponte Costa**[1*]

[1]Computational Neuroscience Unit, Intelligent Systems Labs, Faculty of Engineering, University of Bristol, Bristol, BS8 1TH, United Kingdom; [2]School of Physiology, Pharmacology and Neuroscience, Faculty of Life Sciences, University of Bristol, Bristol, BS8 1TH, United Kingdom

Summary During task execution cortical dynamics must bridge sensory cues with future behavioural outcomes. However, how cortical networks acquire such task-specific dynamics remains unclear. Here we propose that the cerebellum drives cortical dynamics to enable rapid and flexible task acquisition. We model cerebellar networks that are tuned through timing rules to provide cortical networks with task-outcome predictions. First, using sensorimotor tasks we show that cerebellar feedback with fixed cortical connectivity is sufficient for rapid task acquisition and one-shot task switching. Next, we demonstrate that, when trained in working memory tasks, the cerebellum can also underlie the maintenance of cognitive-specific dynamics, explaining a range of optogenetic and behavioural observations. Finally, we use our model to introduce a systems consolidation theory in which task information is gradually transferred from the cerebellum to the cortex. In summary, our results suggest that cortico-cerebellar loops are critical for task acquisition, switching and consolidation in the brain.

Introduction

Learning to interact with the environment requires a continuous integration of fast-changing sensory cues with future behavioural outcomes. Growing evidence suggests that cortical dynamics integrate the task-specific information that is needed for such sensory-behavioural transformations [1–5]. One dominating view in the field assumes that cortical networks are themselves learnt or optimised leading to the rich dynamics required for task performance [6–8]. However, to help ensure that task-encoding remains stable, cortical plasticity must be kept under control and relatively weak [9,10]. This raises the question of how can the cortex acquire new task-specific dynamics in the presence of relatively fixed connectivity?

One possible solution is to consider feedback loops that drive cortical dynamics [11]. Computational studies have extended recurrent neural networks (RNNs) models of cortical networks (Fig. 1a) to incorporate feedback loops for task acquisition. One type of feedback loop drives RNN dynamics by projecting the readout back to the RNN [12–14] (Fig. 1b). Building on this line of work, two recent theoretical studies have suggested that thalamo-cortical feedback can both prepare and control RNN dynamics to achieve flexible motor sequencing [15,16]. All of these studies assume connectivity within the RNN itself remains fixed, thereby avoiding complex learning rules while being able to reuse RNN dynamics for different contexts [17]. However, these approaches either assume a relatively simple feedback (i.e. a linear combination of RNN activity) or rely on theoretically optimal, but biologically implausible , derivations for the feedback signal. In particular, the possible role of more powerful, highly adaptable brain regions are often overlooked.

Here we focus on the feedback loop between two key brain regions, the cortex and *the cerebellum*. The cerebellum is a highly plastic system and is well placed to drive cortical dynamics via a set of stereotypical, but functionally separable cortico-cerebellar loops [18,19]. Indeed, an ever-growing array of clinical [20], functional imaging [21,22], and op-

*For correspondence: [joe.pemberton, rui.costa]@bristol.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

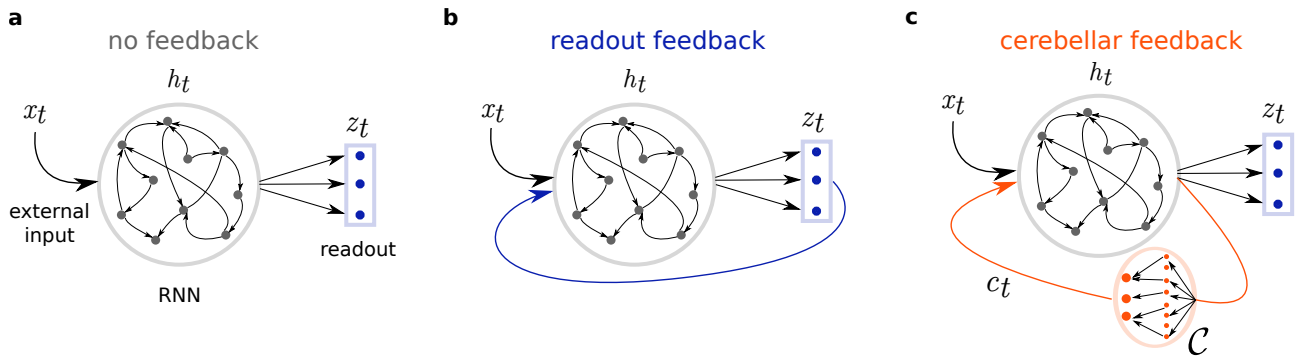

Figure 1. Schematic of cortical recurrent networks with different degrees of feedback. **a**, Model variant with *no feedback:* temporal external input (x_t) is fed to a cortical RNN (grey) and a linear readout layer (blue) produces the final model output (z_t). **b**, Model variant with *readout-only feedback:* in this scheme there is a feedback loop in which the RNN also receives readout predictions as extra input[12,14]. **c**, Model variant with *cerebellar feedback:* an copy of RNN activity (h_t) is sent to a (feedforward) cerebellar network $\mathcal{C}$, which feedbacks to the cortical network its own cerebellar predictions (c_t).

togenetic[23–25] studies support an important cerebellar contribution to cortical activity in both motor and non-motor domains. Recently, two schools of thought have emerged on the computational role of cortico-cerebellar loops. The first asserts that the cerebellum reinforces cortical-dependent goal-directed behaviour by appropriately steering or stabilising cortical states in real-time[26,27]. The second also promotes the cerebellum as a facilitator of goal-directed cortical transitions, but that it does so indirectly via teaching signals which lead to cortical plasticity[28,29]. Whilst these two views may co-exist, it is the former that is well placed to operate under weakly plastic cortical networks. Moreover, the cerebellum acting as an instantaneous driver of cortical dynamics is in line with the fast activity-dependent cortico-cerebellar interactions that have been observed experimentally[23–25].

Here we put forward a computational framework in which the cerebellum learns to rapidly steer and stabilise task-dependent cortical dynamics. We test this model on a variety of motor and non-motor tasks, proposing that the cerebellum is optimised to support task acquisition in the cortex. This reduces the burden of learning in cortical networks and allows a given cortical area to rapidly switch between different tasks. In line with this, we show that a strong cortical dependence on cerebellar feedback arises after learning, consistent with recent behavioural and optogenetic experiments. Finally, we use this model to put forward a cerebellar-to-cortical systems consolidation theory, in which quickly learnt task-specific information encoded by the cerebellum is gradually transferred to the cortex. Overall, we introduce a computationally and experimentally supported theory for cerebellar-supported task acquisition, switching and consolidation in the brain.

Results

A theory of cerebellar-driven cortical dynamics for task acquisition

To study the role that cerebellar feedback can have in driving cortical dynamics during task acquisition, we explore different variants of cortical RNNs: without feedback (Fig. 1a), with readout feedback (Fig. 1b)[14,30] and with feedback provided by a cortico-cerebellar loop (Fig. 1c). We introduce a model of cortico-cerebellar loops, in which a *cortical* RNN is reciprocally connected to a feedforward *cerebellar* network $\mathcal{C}$. In our model, temporal RNN representations $\mathbf{h}_t$ are passed onto the cerebellar network to compute task-specific predictions $\mathbf{c}_t$, which are then sent back to the same cortical RNN. The final model output $\mathbf{z}_t$ is then a linear readout of the RNN activity

$$\begin{aligned} \mathbf{h}_t &= \alpha \mathbf{h}_{t-1} + W_{hh} f(\mathbf{h}_{t-1}) + W_{ih}\mathbf{x}_t + W_{\mathcal{C}h}\mathbf{c}_t \\ \mathbf{z}_t &= W_{\text{rdt}} f(\mathbf{h}_t) \end{aligned} \tag{1}$$

where α denotes the cortical internal memory (or leak) of the RNN neurons, $f(x)$ is the cortical activation function which is set as $\tanh(x)$. W_{hh}, W_{ih}, $W_{\mathcal{C}h}$ are the recurrent, input, and cerebellar weights onto the RNN respectively, and W_{rdt} are the readout weights (see Extended Data Fig. S1 for a detailed schematic). For computational efficiency and due to the relatively long duration of the tasks we train our model using a discrete approximation of a continuous

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

RNN (see Methods). To highlight the need for optimised network connectivity rather than inherent cortical memory mechanisms, in our experiments we generally focus on small $\alpha = 0.1$ (see Methods).

The cerebellar feedback $\mathbf{c}_t$ is a feedforward computation $\mathcal{C}$ on the previous RNN activity

$$\mathbf{c}_t = \mathcal{C}\left(f\left(\mathbf{h}_{t-1}\right)\right) = W_{\text{PF}} f^{\mathcal{C}}\left(W_{\text{MF}} f(\mathbf{h}_{t-1})\right) \tag{2}$$

where W_{MF} represent the cerebellar (input) mossy fibre (MF) weights onto granule cells (GC) and W_{PF} the parallel fibre (PF) weights from GC to Purkinje cells (PC). Together, these constitute the main stages of processing in the cerebellar cortex[31]. In general we model W_{MF} as highly divergent with an input/output ratio of 1:20 (see Methods) and $f^{\mathcal{C}}(x)$ as a rectified linear function (ReLU), in line with the large numbers of cerebellar GCs and their relatively linear processing[32].

We use biologically plausible gradient descent[33] to optimise cortical weights during the acquisition of a given task (Eq. 1). In particular, we minimise the temporal error $E_t = \mathcal{E}(\mathbf{z}_t, \mathbf{y}_t)$, where $\mathbf{y}_t$ denotes the desired task outcome at time t and $\mathcal{E}$ is the task error function (see Methods). These weights can all be optimised simultaneously during learning – we refer to this case as *fully plastic*. However, a key idea that we put forward in this study is that it is not the neocortex, but in fact the cerebellum, which acts as a key driver for task acquisition. For this reason we highlight the case in which RNN plasticity is constrained. In particular, we focus on conditions in which RNN plasticity is either absent – *fixed RNN* case, or in which plasticity is strictly limited to its input synapses (i.e. only W_{ih}, $W_{\mathcal{C}h}$ in Eq. 1 are plastic) – *input plastic* case. The latter case considers both plasticity at sensory and cerebello-cortical input during task acquisition, in line with experimental observations showing plasticity at cerebellar pathways to the cortex[34,35].

In contrast to cortical learning, the cerebellum is always optimised, through a separate but related cerebellar error $E_t^{\mathcal{C}}$ by adjusting its parallel fibre weights W_{PF} (we assume mossy fibre weights to be fixed in line with experimental observations). Like the cortical prediction error the cerebellar error function depends on the desired outcome $\mathbf{y}$. Importantly, to enable the cerebellum to learn to estimate future outcomes ($\mathbf{c}_t \approx \mathbf{y}_{t+\tau}$) we formulate a temporal cerebellar learning rule. In this rule the cerebellum learns by comparing its own past output (within a time-window τ) with future desired outcomes, i.e. $E_t^{\mathcal{C}} = \mathcal{E}(\mathbf{c}_{t-\tau}, \mathbf{y}_t)$. This plasticity rule is consistent with time-dependent plasticity rules observed experimentally at parallel fibre synapses[36]. For our motor-based tasks we generally consider a cerebellar time window of $\tau \approx 150$ms[36] and for the later cognitive tasks use longer windows $\tau \approx 600$ms (see Methods).

Cerebellum learns to drive cortical dynamics during a line drawing task

To study the functional consequences of cortico-cerebellar loops we first test the model in a motor-based line drawing task. In this task the model receives one out of six cues at the beginning of the task and learns to either remain still or produce one out of five possible straight lines (Fig. 2a; see Methods). Feedback provided by desired outcomes (i.e. straight lines) is provided at each timestep. Consistent with behavioural studies on cerebellar patients[37], we find that cerebellar feedback significantly improves learning of the task and final performance (Fig. 2a,b). The ability for cerebellar feedback to facilitate learning does not depend on the degree of plasticity and internal memory in the cortical RNN (Fig. 2c). Interestingly, a fixed RNN with a plastic cerebellum achieves the same learning performance as a fully plastic or input plastic RNN. In contrast, when no feedback or a simple readout feedback is provided the network can fail to learn the task due to fading memory properties of RNNs (Fig. 2b,c). Classical cerebellar models pose that the cerebellum can act as a direct controller of motor tasks[31]. To contrast this view with our model we also train an RNN with a direct cerebellar readout, which apart from the cortico-cerebellar feedback weights uses the same free network parameters, and find it insufficient to learn the task (Extended Data Figs. S1 and S2).

Next, we study how two known cerebellar features: (i) a large number of granule cells and (ii) timed plasticity rules contribute to task proficiency[36]. We find that a combination of high numbers of granule cells with a learning rule with a non-zero temporal horizon, τ, result in better cerebellar learning (Extended Data Fig. S3), which in turn drives better cortical representations and overall task performance (Fig. 2d and Extended Data Figs. S3,S4). Moreover, because both the cortical RNN readout and cerebellar network are trained on the same desired outcome, we observe that cerebellar output effectively predicts cortical readout τms ahead (Fig. 2e). Our model thus provides a theory of how the cerebellum learns to predict upcoming movements[38,39].

The advantage of a large number of granule cells has been well studied is likely due to better linear separability of its inputs[40]. However, what are the computational advantages of the cerebellum providing the cortical RNN with expected future outcomes? Due to fading properties of RNNs sensory cues are rapidly forgotten. Therefore a high

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

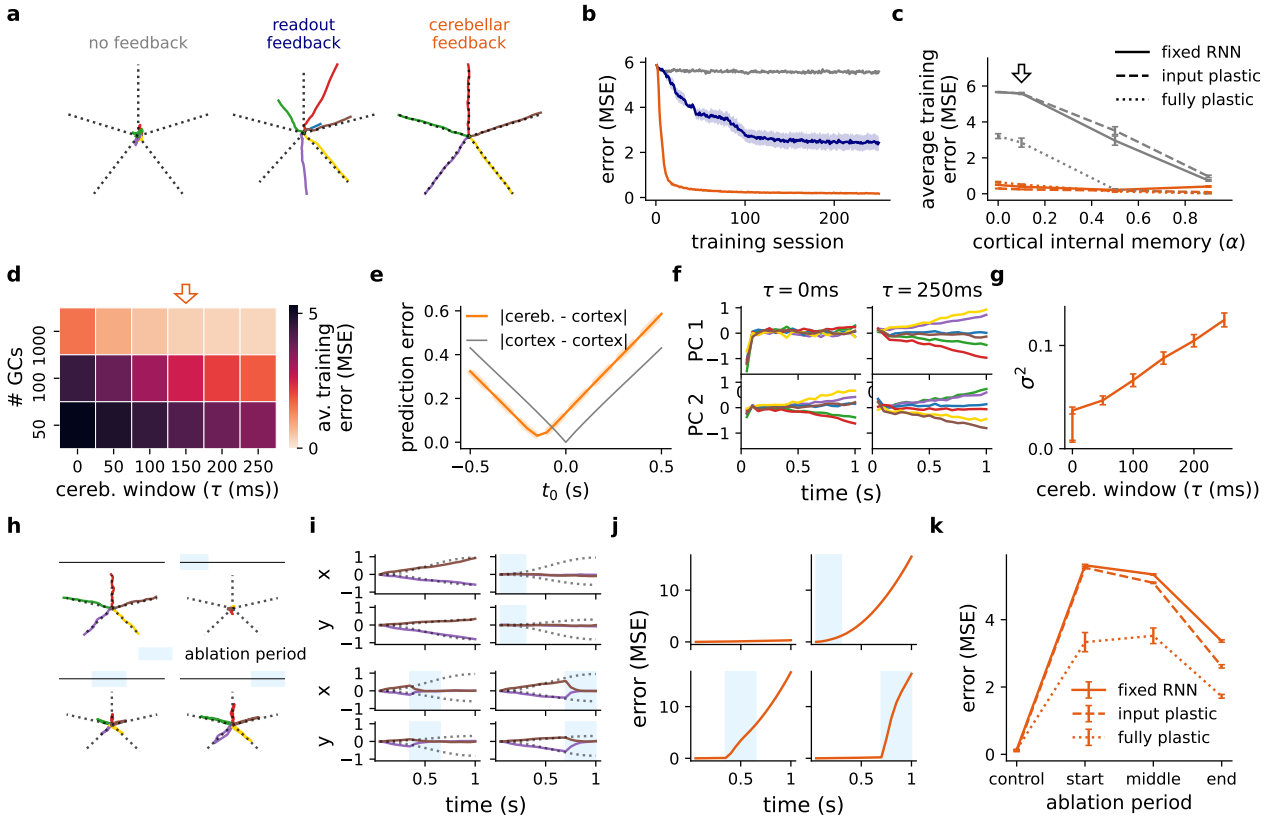

Figure 2. Cerebellum learns to drive cortical dynamics during a line drawing task. **a**, Given one of six possible stimuli at the first timestep the model must learn to draw a corresponding line (dotted black line) or remain still. Model output after training is shown for three model architectures with a fixed RNN. **b**, Learning curves of models in A (same colour-coding). MSE denotes mean squared error. **c**, Average training error across different levels of RNN internal memory (α) and plasticity (fixed RNN, input plastic and fully plastic) for the no feedback and cerebellar feedback models; arrow denotes cortical internal memory used in the other panels ($\alpha = 0.1$). **d**, Average training error of cortico-cerebellar model under varying numbers of granule cells and cerebellar temporal windows (τ). Orange arrow denotes default parameter choices. **e**, Prediction error between cortical output and itself (gray) or cortical output and cerebellar output (orange) for different temporal delays. **f**, Evolution of first (upper panel) and second (lower panel) principal components of cortical RNN for different stimuli, colour-coded as in **a**, using small ($\tau = 0$ms) and large ($\tau = 250$ms) cerebellar time windows. **g**, Average variance of first two principal components across different cerebellar temporal windows, τ. **h**, Model output for different periods of cerebellar ablation (blue box represents period of ablation). **i**, Output x and y coordinates of the lines drawn in **h**. **j**, Average model error across all inputs for ablation periods in **h**,**i**. **k**, Average error for different degrees of plasticity and ablation periods (left to right) as in **h**-**j**. All results are averaged over 5 seeds. Error bars represent standard error of the mean.

cerebellar τ gives the cerebellar network the ability to map RNN activity to desired outcomes early on in the task. Consistent with this we find that the predictive cerebellar output drives outcome-dependent RNN representations (Fig. 2f,g). This result showing potent initial drive of cortical activity could provide a justification for the observed role of the cerebellum in movement initiation[41,42].

Finally, to directly examine the role of cerebellar feedback on cortical dynamics, we inhibit - or "ablate" - cerebellar output (i.e we set $\mathbf{c}_t = 0$ in Eq. 1) during different stages of the task. In each case we observe significant impairment in the model output which returns to baseline (Fig. 2h-j). Moreover, this effect is most detrimental to task performance when ablation occurs at the start (Fig. 2k). These findings are consistent with the observed freezing effect of cerebellar lesions on gait[43]. In line with both cortical and cerebellar networks working jointly to perform the task, we find that when the RNN is fully plastic cerebellar ablations have a significant but reduced impact on the cortical dynamics (Fig. 2k and Extended Data Fig. S5). We also observe that the cortical RNN is particularly sensitive to the presence of noise in cerebellar output. When noise is added to its output it leads to irregular behaviour (Fig. S6), in line with the classical motor symptoms of cerebellar ataxia[44].

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Taken together, this motor-based task highlights the computational benefits of training a cerebellar network to drive cortical dynamics, predicting that the cortex can critically depend on cerebellar feedback for successful task execution. Furthermore, we demonstrate that cerebellar plasticity can effectively replace the need for local cortical plasticity.

Cerebellar-mediated task switching in cortical networks

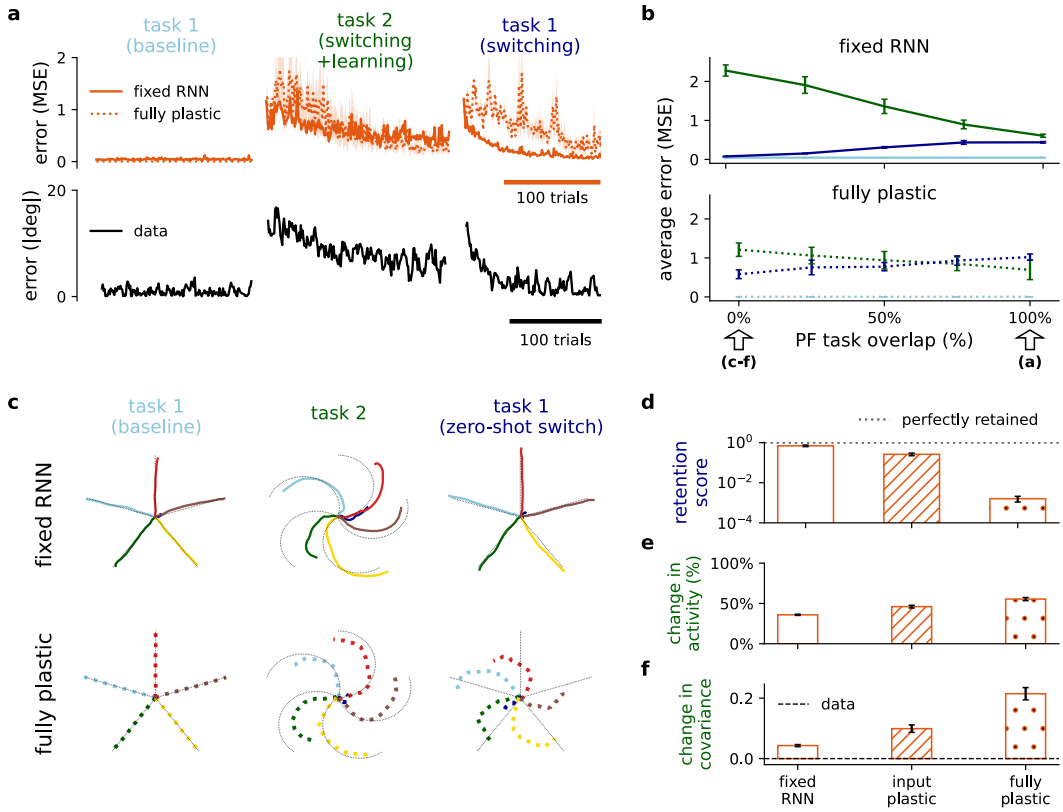

Figure 3. Context-dependent cerebellar feedback can enable multi-task learning and switching in the cortex. **a**, Training error of cortico-cerebellar models originally trained for line drawing (cf. Fig. 2; $\alpha = 0.5$). The models continue to execute the line-drawing task (left) before being trained on a novel curl-field variant of the task (middle) and then finally switch back to the original task (right). Data from behavioural experiments in macaque monkeys is reproduced here for comparison (bottom; [45]). **b**, Average training error across different levels of parallel fibre (PF) task overlap for the different tasks for the fixed RNN (top) and fully plastic (bottom) models. Task periods colour-coded as in **a**. Arrows denote degree of PF task overlap highlighted in **a** and **c**-**f**. **c**, Model output for each of the three training periods defined in A for the zero-overlap condition; "zero-shot" output corresponds to the model output in the first trial when task 1 is reintroduced. **d**, Model retention score for task 1. The retention score is computed as the error of task 1 during baseline over the error at the first trial after switching back to task 1. **e, f** Change in **e** activity and **f** covariance in the RNN population between task 1 (baseline) and after learning task 2. Mean changes in experimental data in **f** are reproduced (see Methods) from neuronal recordings obtained from premotor (PMd) and primary motor (M1) cortices in macaque monkeys [45]. All results are averaged over 5 seeds. Error bars represent standard error of the mean.

We have shown that cortico-cerebellar loops can enable successful task learning with minimal cortical plasticity. This opens the possibility of reusing cortical networks across different contexts and behaviours.

To demonstrate the model's ability to adapt and perform context-dependent task switching, we consider how models trained in the line-drawing task can be retrained to a curl-field variant [45]. In particular, we analyse how the cerebellar network can (i) successfully enable learning in a new task context and also (ii) rapidly revert, or *switch*, to a previously learned context.

As expected, when the new task context is introduced to the model, there is a steep increase in error before the model successfully learns the new task (Fig. 3a, left and middle). Notably, however, when the original task is reintroduced, the fixed RNN model recovers the initial dynamics significantly faster than the fully plastic model and

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

more faithfully captures the behavioural data from macaque monkeys[45] (Fig. 3a, right). This relatively slow switching back suggests that the fully plastic RNN is more prone to forgetting the original task[9].

We then asked how the cerebellar network might enable even faster task switching. In line with observed context-dependent activations[46] and plasticity rules[47] in the cerebellum, we consider cerebellar PFs which are task-specific. The extent of task-specificity at PFs is modelled by the *PF task overlap*; full overlap (100%) would imply that the same exact PFs are used across task contexts, while zero overlap (0%) implies that a completely different set of PFs is used for each task respectively.

Our results show that the degree of PF task overlap predicts a tradeoff between the speed of learning the new task and the ability to rapidly switch back to the original task (Fig. 3b). Specifically, whilst maximal PF task overlap is beneficial when a new task is introduced, rapid switching is favoured when distinct PFs are used. To highlight the ability to immediately switch back to the original task (zero-shot switch) we focus on the zero-overlap case. For the fixed RNN, but not the fully plastic RNN, the model achieves near-perfect switching to the original task (Fig. 3c,d). Consistent with the need to learn a new task all models show a substantial change in the neuronal activity (Fig. 3e and Extended Data Fig.S7a). However, we expect that models with minimal local cortical plasticity should result in minimal changes in the underlying dynamics of both tasks. To test this, we measure changes in the the covariance of the neuronal activity between the new task and the initial task (see Methods and[48]). As predicted, only the models with reduced cortical plasticity show the minimal changes observed experimentally (Fig. 3f and Extended Data Fig. S7b). On the other hand, the fully plastic model the dynamics acquired after switching back to the initial task are significantly different to baseline (Extended Data Fig. S7c,d). This suggests that the fully plastic model learns a new solution to the initial task, explaining its relative slowness in switching.

Overall, we apply our models to demonstrate a cerebellar-driven solution to multi-task learning and task switching. We show that the underlying dynamics preserved by a fixed cortical RNN, supported by context-dependent cerebellar feedback, can support rapid behavioural changes whilst minimising forgetting of previously acquired task knowledge.

Cerebellar temporal basis supports non-linear drawing task

Above we have modelled a case in which the cerebellum learns to drive cortical dynamics using a specific predictive time-window (namely $\tau = 150$ms). However, a recent study has revealed a diversity of temporal plasticity windows to be at play in the cerebellum[36,49] (Fig. 4a). Such diversity of temporal windows may enable the cerebellum to learn a *temporal basis* for upcoming events, which may enhance the cerebellum's ability to predict future outcomes.

To demonstrate the benefit of diversity in temporal windows we consider a more realistic (and challenging) variant of the line-drawing task in which the model is now trained to produce a digit-like output (Fig. 4b; see Methods). This task is selected so as to produce a non-linear and highly varied set of future desired outcomes and therefore the need for richer cerebellar predictions. In particular, we consider a cerebellar network which simultaneously learns with a range, or "temporal basis", of time-windows $\tau \in [0\text{ms}, 250\text{ms}]$ such that its prediction effectively spans a relatively long window of upcoming desired outcomes (see Methods).

We find this heterogeneity of cerebellar time windows to enable both faster learning and higher performance thresholds (Fig. 4b,c and Extended Data Fig. S8). As expected, when considering the simpler line-drawing task having multiple time windows does not improve learning (Fig. S8c). Moreover, in line with the results above, a fixed RNN achieves a performance comparable to the plastic RNN models across different degrees of internal memory in the cortical network (Fig. 4d). When comparing the network performance across different numbers of granule cells and time-windows, we find that higher numbers of granule cells combined with multiple time-window learning achieves the best average learning performance (Fig. 4e). Finally, as with the simpler line-drawing task, we find that cerebellar ablation is detrimental to the maintenance and development of these representations (Fig. 4f-h and Extended Data Fig. S9).

These results suggest that the diversity of cerebellar temporal windows observed experimentally[36,49] can be beneficial when in the presence of more challenging task conditions.

Cerebellar-driven cortical dynamics maintains beliefs in an evidence accumulation task

So far we have focused purely on motor-based tasks, but growing evidence strongly suggests that the cerebellum also plays important roles in functions that go beyond direct motor control[19,50]. To demonstrate this we model an evidence accumulation task that has been shown to be cerebellar-dependent[24]. In this study Deverett et al.[24] showed

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

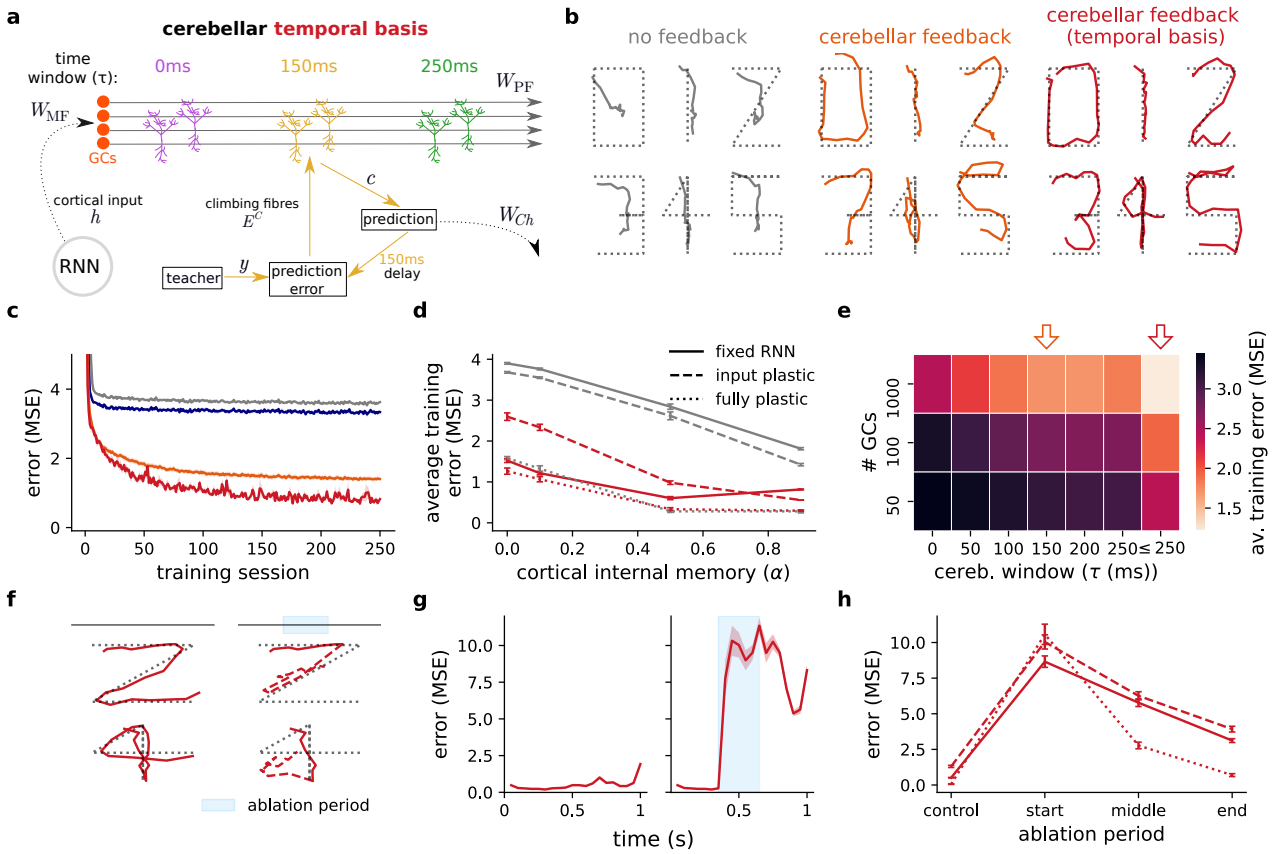

Figure 4. Cerebellar temporal basis supports cortical dynamics of a non-linear digit drawing task. **a**, Schematic of cerebellar learning with a temporal basis. We consider multiple populations of Purkinje cells with different learning time windows τ. **b**, Model output after training for different input examples of the digit drawing task (fixed RNN; $\alpha = 0.1$). **c**, Learning curves of models in **b** together with readout feedback model (blue). **d**, Average training error across different levels of RNN cortical internal memory (α) and plasticity assumptions. **e**, Performance of cerebellar feedback for different numbers of granule cells and and cerebellar time windows. Orange arrow indicates default parameter choices with a single cerebellar time window; red arrow indicates temporal basis model with multiple time-windows. **f**, Model output under control and cerebellar ablation conditions for example inputs (digit 2 in upper panels and digit 4 in lower panels); dashed red line represents model output during and after ablation period. **g**, Average model error across all inputs for control (left) and ablation (right) conditions. **h**, Average error for different degrees of cortical plasticity and ablation periods (middle period illustrated in (F,G)). All results are averaged over 5 seeds. Error bars represent standard error of the mean.

that optogenetic inhibition of the cerebellar output nuclei disrupts the ability of mice to determine whether the left or right cheek received more air puffs over a period of time (Fig. 5a). Unlike the previous tasks, here the desired outcome is only provided at the end of the task, making error-related signals highly sparse.

Similar to the motor tasks studied above, cerebellar feedback improves task learning relative to models without feedback or with readout feedback (Fig. 5b). Moreover, a fixed RNN achieves performance comparable or even superior to the fully plastic models across a range of degrees of cortical internal memory (Fig. 5c and Extended Data Fig. S10). These results suggest that weakly plastic cortical networks driven by the cerebellum may also be sufficient for learning cognitive-based tasks with sparse error information.

Next, our ablation analysis reveals strong similarities to the optogenetic observations by Deverett et al.[24]. In particular, cerebellar ablation greatly impairs the model's capacity to maintain and develop beliefs, mirroring the behavioural effects observed experimentally (Fig. 5d and Extended Data Fig. S11). The general trend observed experimentally – that later cerebellar ablation is particularly detrimental to task performance – is also captured by our model (Fig. 5e). Indeed, using the same behavioural regression performed by Deverett et al.[24] (see Methods), we show that cerebellar ablation in latter periods leads to a final choice in which information about previously seen inputs is greatly reduced (Fig. 5f), in line with experimental findings. We further quantified this inability to integrate prior

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

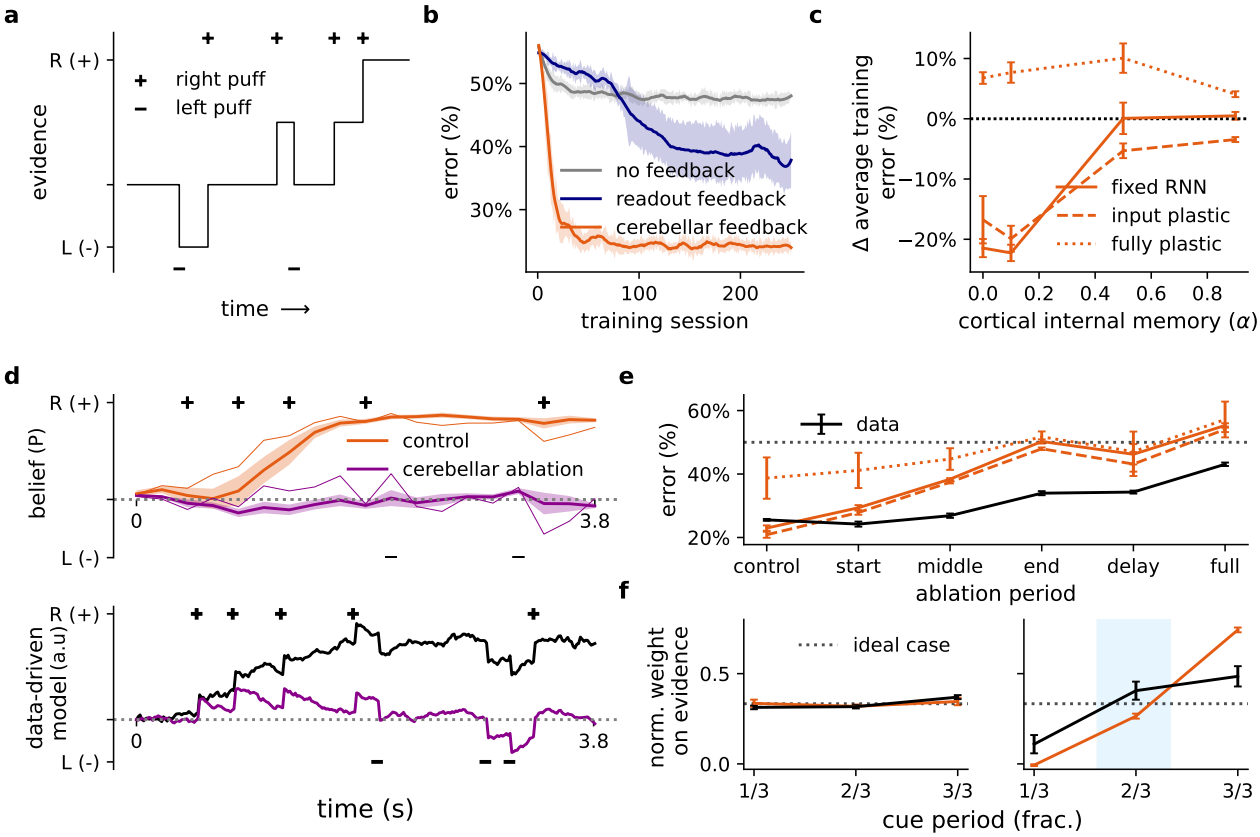

Figure 5. Cortico-cerebellar model mimics mouse behaviour during evidence accumulation task. **a**, Schematic of evidence accumulation task[24]: a random sequence of non-zero inputs ("air puffs") is delivered in the leftward (-) or rightward (+) direction. The model must integrate this input and decide at the end of the task which side received more input overall. **b**, Learning curves of models (fixed RNN; $\alpha = 0.1$) without feedback (grey), with readout feedback (blue) and with cerebellar feedback (orange). **c**, Change in average training error of the cortico-cerebellar model with respect to the no feedback model across different levels of cortical internal memory (α) and degrees of cortical plasticity (see Extended Data Fig. S10b). **d**, Model beliefs over time without (orange) and with complete cerebellar ablation (purple) in model (upper panels) and data-derived behavioural model (lower panels) reproduced from Deverett et al.[24]. Thin model lines represent one example seed. Belief P denotes model output probability. **e**, Model and data error under different ablation periods and degrees of cortical plasticity. **f**, Normalised regression weights at different periods of input presentation (cue) during control (left) and ablation (right) conditions for both model (orange line) and behavioural data (black line). All model results are averaged over 5 seeds. Error bars represent standard error of the mean.

input by analysing performance specifically for "history-centric" trials, which rely on remembering the history of input in the first third period of the cue presentation (see Methods). As expected, late cerebellar ablation is particularly detrimental for these trials (Extended Data Fig. S11).

Overall, our model predicts that the proper maintenance of model selectivity depends critically on cerebellar feedback during evidence accumulation. Consistent with behavioural results, these effects are emphasised when cerebellar ablation occurs in the later stages of the task.

Cerebellar feedback sustains cortical dynamics in a delayed association task

Next we aim to demonstrate that cerebellar networks can also effectively drive cortical dynamics in tasks with long delay periods, while capturing both neuronal and behavioural observations. To achieve this we model a delayed association task which was recently shown to dependent on cortico-cerebellar loops[23]. In this study mice were presented with one of two stimuli (left or right) followed by a delay period, after which they were trained to lick in the corresponding direction (Fig. 6a, top). At the same time neural selectivity was recorded both in the anterior lateral motor cortex (ALM) - a working memory and planning region - as well as the cerebellar output nuclei (Fig. 6a, bottom). Timed photoinhibition was used to reveal ALM selectivity to strongly depend on the cerebellar output nuclei, and vice

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

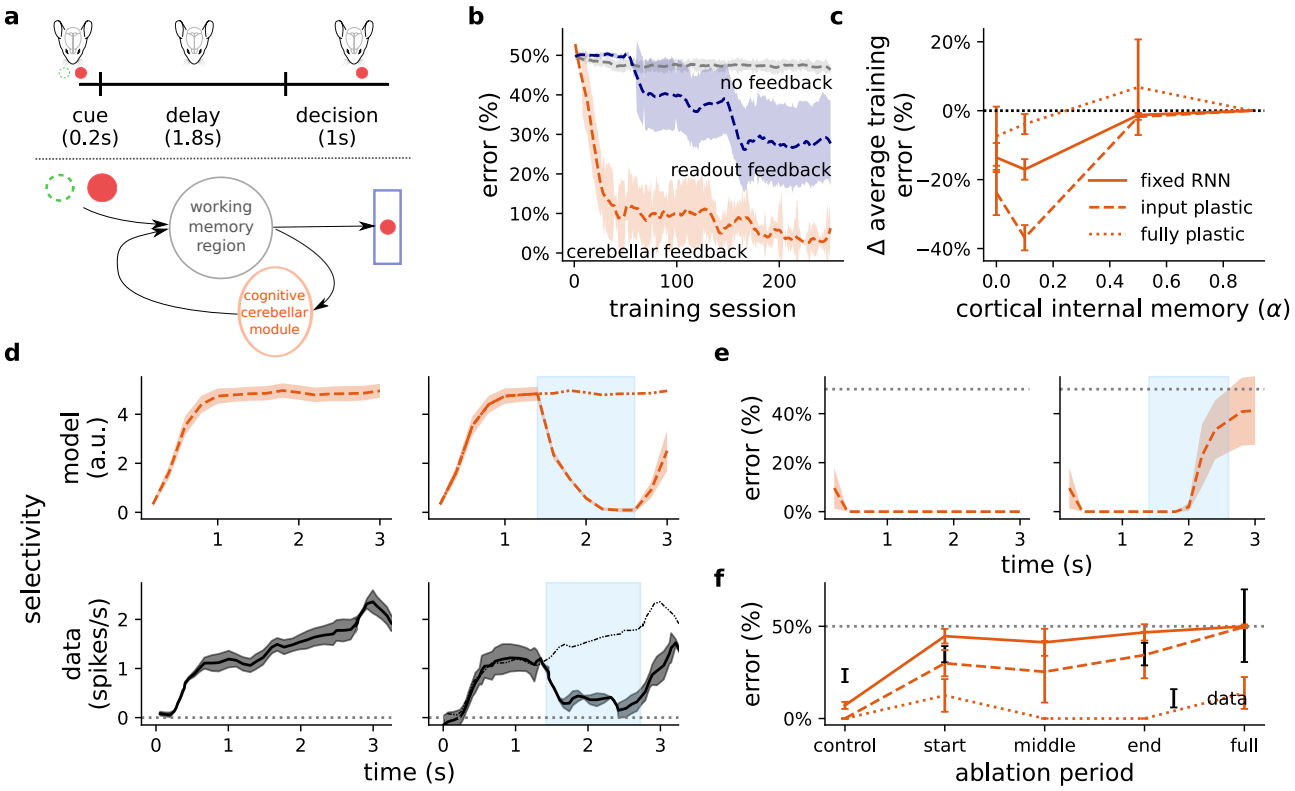

Figure 6. Cerebellar network sustains cortical dynamics during delayed association task in line with optogenetic experiments. **a**, (Top) Delayed association task; a sensory cue is presented followed by a delay and decision period[23]. (Bottom) The cortico-cerebellar loop models the interactions between a working memory region and a cognitive module of the cerebellum. **b**, Learning curves of model without feedback (grey), readout feedback (blue) or cerebellar feedback (orange) for models with an input plastic RNN ($\alpha = 0.1$) . **c**, Change in average training error of the cortico-cerebellar model with respect to the no feedback model across different levels of cortical internal memory (α) and degrees of plasticity in the cortical RNN (see Extended Data Fig. S12b). **d**, Cue selectivity during the delay period without (left) and with cerebellar ablation (right; blue area denotes period of ablation and thin line shows control) in the model (upper panels) and optogenetic experiments (lower panels) reproduced from Gao et al. [23]. **e**, Model error during the delay period without (left) and with (right) cerebellar ablation. Dotted line denotes chance level. **f**, Average error from cerebellar ablation at different points during the delay period and different degrees of cortical plasticity. Experimental data is shown in black. Results are shown across 5 seeds. Error bars represent standard error of the mean.

versa.

To model this task we follow the same protocol used experimentally[23], where one of two possible cues are presented followed by a delay period, after which the model makes a cue-based response (left or right; see Methods). Given the lack of sensory or teaching information during the delay period the cortico-cerebellar network it is particularly vital in this task to sustain stimulus representations. It is important to note that a standard randomly initialised RNN is unlikely to achieve this property, since memories of previous inputs naturally fade[17] in the absence of task-induced plasticity.

We observe that cerebellar feedback consistently enables task acquisition (Extended Data Fig. S12), and identify a particularly interesting case when plasticity in the RNN is limited strictly to its input synapses (input plastic). In this case cerebellar feedback significantly improves cortical learning to reach near-perfect performance, whilst also enabling a high degree of stability in task selectivity throughout the delay period (Fig. 6b-d and Extended Data Fig. S12). We speculated that for this task input plasticity is particularly important, because the cerebellum is required to sustain task-specific predictions in the RNN throughout the entire delay period. We verified this stronger cerebello-cortical drive by using concepts from control theory[51]. In particular, we can explicitly relate cerebello-cortical optimisation to a quantitative increase in the impact, or *energy*, of cerebellar feedback onto RNN activity (Extended Data Fig. S13; see Methods). Moreover, the ability of the cerebellum to drive cortical dynamics should depend on the cortical network's ability to express those dynamics. In line with this view our results show that (even untrained) cortical recurrent

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

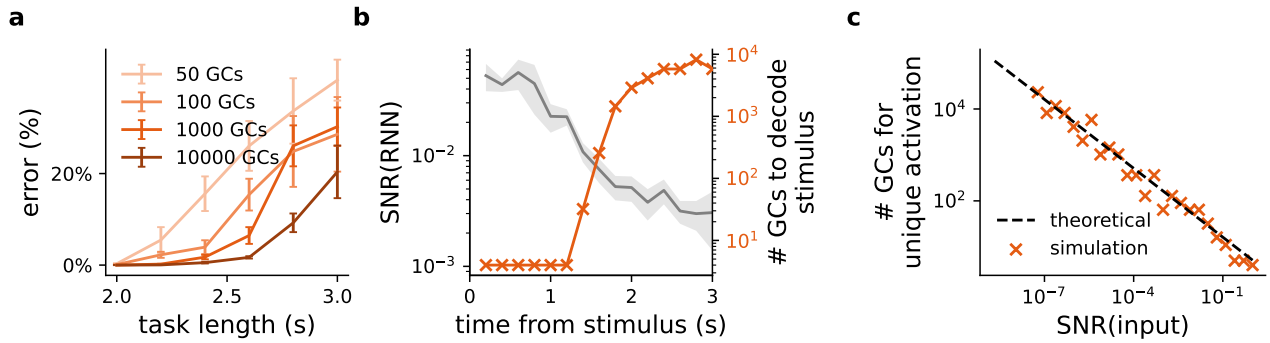

Figure 7. A large granular layer enables the cerebellum to decode fading cortical memories. **a**, Model error for different numbers of cerebellar granule cells (GCs) and delay period lengths in the delayed association task (fixed RNN; $\alpha = 0.1$). **b**, Signal-to-noise ratio (SNR) of RNN activities (left y-axis) and number of GCs needed to decode the stimulus from these activities (right y-axis). **c**, Number of GCs needed for distinct granular population responses for different SNRs using synthetic inputs; dashed black line is theoretical prediction (see main text), markers are results from numerical simulation. Results in **a** are averaged over 5 seeds; results in **b**,**c** are averaged over 10 seeds. Error bars represent standard error of the mean.

weights are important in maintaining cerebellar predictions over time (Fig. S14).

Next, to demonstrate that the cerebellum helps drive task-specific dynamics in the cortical RNN we performed a simulated ablation in which the cerebellum is transiently removed during the delay period. We find that cerebellar ablation drastically disrupts cortical task selectivity, whilst cerebellar selectivity is also highly disrupted by cortical ablation (Fig. 6d-f and Extended Data Fig. S15). These ablation results are consistent with changes in cortical dynamics observed during optogenetic-mediated ablation[23] (Fig. 6d, bottom and Extended Data Fig. S15a, bottom). Our model predicts that this effect should depend on the degree of plasticity in the cortical RNN. Consistent with this view in the presence of a fully plastic RNN we do not observe the strong cerebellar dependence that was observed experimentally (Fig. 6f and Extended Data Fig. S16). Taken together our results suggest that the cerebellum, not the neocortex, is the primary site of learning during the acquisition of this working memory task[23].

Overall, these results demonstrate that our model can capture working memory tasks and the observed dependency of cortical dynamics on cerebellar input. Moreover, our model makes the prediction that the cerebellum is a key site of plasticity during acquisition of delayed association tasks.

Computational power of cerebellum can decode fading cortical memories

A prevalent feature in classical cerebellar theories is that the divergence provided by the granular layer enables better separation of its inputs[52]. This is particularly important for cases in which input activity shows little variance across conditions and the original task-related signal is only weakly encoded, for which a large dimensional expansion at the granular cell layer would help to recover the original task-related signal[53].

In the delayed association task, for which memory of the initial stimulus quickly fades with time in the untrained RNN, we observe that as the delay period increases so does the need for more granule cells (Fig. 7a). Indeed, by considering the cerebellar network as a direct readout of the RNN, we observe an inverse relationship between the signal-to-noise ratio (SNR) of the RNN activity and the required number of granule cells to decode the signal (Fig. 7b). We formalise this relationship analytically, and show under reasonable assumptions that the required number of granule cells for distinct stimulus encodings is approximately $\frac{C}{\sqrt{\text{SNR}}}$, where $C = -\frac{\pi}{\sqrt{2}}\log(1-\theta)$ and $\theta \in [0, 1)$ denotes the level of confidence (see Methods). We find this relationship consistent with our simulations (Fig. 7c) and (as expected) that these distinct population responses enable successful downstream discrimination of the initial stimulus (Fig. S17). With the cerebellar granular layer consisting of vast numbers of cells – > 50 billion in the human brain[54] – receiving distinct combinations of input, it is therefore well placed to decode the potentially rapidly fading memories in cortical networks. This suggests a key cerebellar role in decoding overlapping, long-lasting temporal cortical dynamics, and supports experimental findings showing that cerebellar inhibition is more detrimental for learning when in the presence of long delay periods[55].

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Cerebellar task knowledge can be consolidated in the cortex

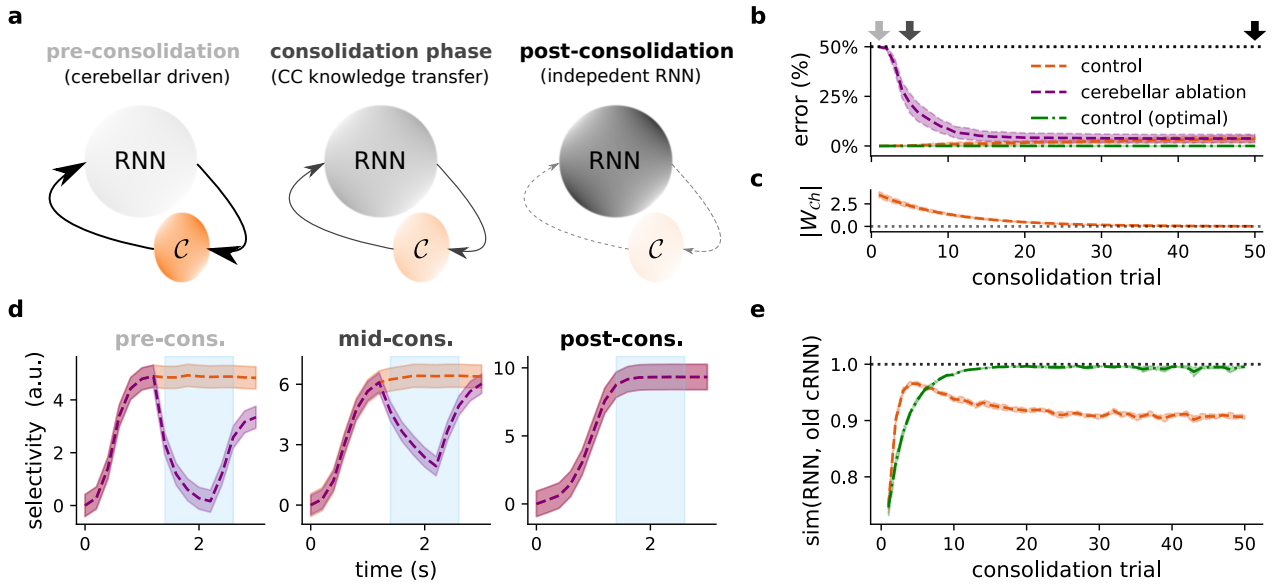

Figure 8. Cerebellum can mediate task consolidation in the cortex. **a**, Schematic of proposed theory of cerebellar-to-cortical task consolidation. During the initial learning phase (left), task representations are primarily driven by the cerebellum and RNN connectivity is not yet task-specialised. During the consolidation phase there is a period of cerebellar-to-cortical (CC) task information transfer (middle), whereby CC interaction drives plasticity in the cortical RNN. After consolidation (right), the RNN can operate effectively without the need for cerebellar input. The colour of the structures reflects the importance of each component throughout consolidation. **b**, Model accuracy in the delayed association task (Fig. 6) throughout consolidation with (purple) and without (orange) cerebellar ablation. For reference an optimal consolidation model is also given (green). Dotted black line denotes chance. **c**, Strength of the cerebellar-to-cortical pathway (W_{Ch}) over the period of consolidation. **d**, Model selectivity with and without cerebellar ablation at different stages of the consolidation process; titles colour coded according to arrows in **b**. **e**, Cosine similarity between cortico-cortical input and total cortical input (i.e. cerebellar-cortical and cortico-cortical inputs) pre-consolidation. Similarity of the consolidation model is shown in orange and the optimal consolidation model in green. Results are averaged over 5 seeds. Error bars represent standard error of the mean.

In each of the previous tasks, cerebellar feedback is shown to mediate learning and the maintenance of task-specific cortical dynamics. However, the neocortex is known to encode long-term representations of tasks[56]. This suggests a need for a "consolidation" period, during which the memory stored in the cerebellum may be transferred to cortical areas.

To demonstrate cerebellar-to-cortical systems consolidation in our model we develop consolidation-specific learning rules. To achieve consolidation we train cortical recurrent weights to mimic cerebellar input (see Methods). Given that (i) the RNN state **h** is of significantly higher dimensionality than the cerebellar output **c**, and that (ii) **c** is a function of **h** (Eq. 2), several solutions that enable the RNN to capture the cerebellar input should exist. We also gradually decay the cerebellar-to-cortical input weights so that over training the cerebellum stops driving the cortical network, thereby giving full control of the task to the cortical RNN (Fig. 8a).

We tested this computational theory of consolidation on the cortico-cerebellar models (input plastic condition) trained on the previous delayed association task (Fig. 6d, top left). We consider two types of learning rule: (i) a simple biologically plausible rule, which depends on the ratio of cerebellar-to-cortical input and total RNN activity and (ii) a theoretically optimal rule based on a least squares solution (see Methods). In both cases, we observe that the RNNs gradually learn to perform the task without the need for cerebellar input (Fig. 8b-d). Indeed, as expected, we observe that the newly learned recurrent weights faithfully replicate the cerebellar input of the original network. (Fig. 8e). In addition, we find that a model with fixed RNN connectivity does not perform as well as the input plastic condition (Extended Data Fig. S18). This is likely due to better network stability when in the presence of the input plastic, compared to the purely fixed RNN (Extended Data Fig. S16a,b).

In summary, the framework we introduce here suggest that the cortico-cerebellar loops may play an important role in systems consolidation by gradually transferring the rapidly learnt cerebellar knowledge to the cortex.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Discussion

Growing experimental evidence suggests that cortico-cerebellar loops support behaviour, but their computational roles have remained unclear. Here we have introduced a systems-level modelling framework in which a feedforward cerebellar network receives the state of a cortical RNN and provides task-specific predictions in return. In our model, cerebellar feedback facilitates learning by shaping the underlying cortical dynamics during motor and cognitive tasks in a way that is consistent with both behavioural and optogenetic studies. Our work suggests that the cerebellum is a key site of learning in the brain, allowing for rapid context-switching of cortical dynamics that underlie behaviour. We finish by introducing a theory of cerebellar-to-cortical system consolidation, in which task-specific knowledge is gradually transferred to the cortical network.

Our model is related to previous network architectures in that it uses feedback to enhance neuronal representations and selectivity in a otherwise fixed RNN, thereby facilitating task-relevant downstream processes[13,14]. However, in contrast with previous work in our model feedback does not stem directly from the (cortical) readout, but from a large cerebellar network with time-predictive learning rules (Fig. 1). The resulting architecture suggests a symbiotic cortico-cerebellar learning strategy: as cerebellar representations improve through rapid learning so do cortical representations, and as cortical representations improve cerebellar decoding becomes easier. According to our model the cortical RNN relies on cerebellar feedback to maintain appropriate representations. Therefore, one can consider the cerebellum as a controller of cortical dynamics. Indeed, two recent theoretical studies have shown how the thalamus, the intermediary in cerebellar-to-cortical communication, can optimally control cortical dynamics during motor preparation and execution[15,16]. In contrast to existing theories, our work suggests that cerebellar-thalamic-cortical pathways do not have the ability, or indeed need, to provide highly precise control of cortical dynamics. Instead our work suggests that the cerebellum can only provide the cortex with relatively general predictions which can be reused across the cortex.

By retraining cortico-cerebellar networks in a novel task we propose a key role of the cerebellum in task switching (Fig. 3). In particular, we show that cerebellar feedback may provide a solution to the problem of context-dependent adaptation, which requires (i) an ability to learn a new context but also (ii) an instant retrieval of appropriate response to previously learned contexts[57,58]. Interestingly, we observe that while recurrent cortical plasticity enables adaptation to a new task context there is catastrophic forgetting of the original context. This is at odds with well-known behaviour in the primate, and provides a computational explanation for why cortical modifications in the monkey cortex during motor adaptation are limited[45].

A unifying model of the cortico-cerebellar loop, and indeed the cerebellum itself, must extend to non-motor tasks. Recent task-based fMRI studies have revealed functional diversity of the cerebellar cortex across a range of cognitive functions[21]. Our model inherently implies a high degree of heterogeneity – it suggests that different modules would be required to drive different parts of the cortex that in turn underlie different cognitive functions. In this study we modelled recent behavioural and optogenetic experimental observations[23,24] which directly implicate the cerebellum in supporting cortical dynamics during evidence accumulation and delayed association tasks (Figs. 5 and 6). In particular, our results show that cortico-cerebellar interactions are enough to learn tasks with highly sparse teaching signals (i.e. only at the end of the task). By combining simulations and theoretical analysis we show that this is enabled by the sheer computational power of the cerebellum (Fig. 7). Moreover, we can best capture experimental observations in conditions in which RNN plasticity is limited, making the prediction that the cerebellum is the primary site of learning for these tasks. This provides an alternative to the commonly assumed view that cortical areas are optimised for specific tasks[6–8].

In our model the cerebellum drives cortical dynamics based on prediction error signals that depend on the desired task outcome. In the case of the working memory tasks and in line with the experimental task setup, the desired outcome can be interpreted as a reward signal. Therefore, from this perspective, the cerebellum learns to predict future rewarding events. This is consistent with the growing literature showing that the cerebellum encodes reward-related signals[59]. However, it remains to be tested exactly how the reward-predictive representations developed by our model compare to those found experimentally.

Here we have also introduced a theory of cerebello-cortical task consolidation. Our theory suggest that cerebellar and cortical learning may operate at different timescales: after an initial fast stage of learning driven by the cerebellum, a period of consolidation might ensue in which the cortex gradually acquires task-specific knowledge encoded

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

in the cerebellum (Fig. 8). Such combination of fast and gradual learning is reminiscent of recent experimental results which suggest significantly faster timescales of plasticity in the hippocampus compared to the prefrontal cortex during a cognitive task[60]. Moreover, the consolidation period can be related to the idea that a task-optimised cerebellum can be utilised as a cortical teacher[28,29]. It is in principle possible for cerebellar-thalamo-cortical projections to support this dual role of the cerebellum as both a driver and teacher of cortical states. Indeed, anatomical evidence suggests that this could occur by providing "driving" and "teaching" input to basal and apical dendrites of cortical pyramidal cells, respectively[61].

Our work highlights commonalities of cortico-cerebellar interactions in motor and cognitive tasks alike. However, it also suggests interesting differences. The first marked distinction relates to the increased significance of cerebellar-to-cortical (input) plasticity during pure working memory (Fig. 6). This is in line with recent experimental evidence showing stronger plasticity at higher-order thalamo-cortical pathways[35]. Indeed, because of the need to sustain information during the delay period without sensory or teaching input, it is advantageous for the network to encode a point attractor-like state (see Extended Data Fig. S16, left). Cerebello-cortical plasticity[34,35] may thus enable greater *controllability* of cerebellar feedback to push the network to these states during working memory tasks, but less so in motor-based tasks[51] (Fig. S13).

Related to the point above, the second difference we highlight is about cerebello-cortical consolidation being more readily achieved when in the presence of networks with stable dynamics (cf. Fig. 8 and Extended Data Fig. S18). We speculate that unstable network dynamics make cerebellar-to-cortical consolidation less stable. Therefore, we predict that while cerebellar-to-cortical systems consolidation might be possible for tasks which involve discrete stable representations (e.g. working memory tasks), for tasks which involve faster, more dynamic responses (as often required in the motor domain) cerebellar control is likely to be required throughout life.

Overall, our work provides a theoretical framework with which to study the function and mechanisms of cerebellar-driven cortical dynamics for rapid task acquisition, switching and consolidation.

Acknowledgements

We would like to thank the Neural & Machine Learning group, James M Shine, Paul Dodson, Everton Agnes, Laureline Logiaco and Jake Stroud for useful feedback. J.P. was funded by a EPSRC Doctoral Training Partnership award (EP/R513179/1), P.C. by the Wellcome Trust (209453/Z/17/Z) and R.P.C. by the Medical Research Council (MR/X006107/1). This work made use of the HPC system Blue Pebble at the University of Bristol, UK.

Author contributions

J.P. developed computational framework with guidance from R.P.C. J.P. performed all numerical and analytical work. J.P. and R.P.C. wrote the manuscript, with contributions from P.C. R.P.C supervised the project.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

[1] Wael F Asaad, Gregor Rainer, and Earl K Miller. Task-specific neural activity in the primate prefrontal cortex. *Journal of neurophysiology*, 84(1):451–459, 2000.

[2] Earl K Miller and Jonathan D Cohen. An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, 24(1): 167–202, 2001.

[3] Jun Tanji and Eiji Hoshi. Role of the lateral prefrontal cortex in executive behavioral control. *Physiological reviews*, 88(1):37–57, 2008.

[4] Farshad A Mansouri, Keiji Tanaka, and Mark J Buckley. Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nature Reviews Neuroscience*, 10(2):141–152, 2009.

[5] Abhishek Banerjee, Giuseppe Parente, Jasper Teutsch, Christopher Lewis, Fabian F Voigt, and Fritjof Helmchen. Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature*, 585(7824):245–250, 2020.

[6] H Francis Song, Guangyu R Yang, and Xiao-Jing Wang. Reward-based training of recurrent neural networks for cognitive and value-based tasks. *Elife*, 6:e21492, 2017.

[7] A Emin Orhan and Wei Ji Ma. A diverse range of factors affect the nature of neural representations underlying short-term memory. *Nature neuroscience*, 22(2):275–283, 2019.

[8] Mikio C Aoi, Valerio Mante, and Jonathan W Pillow. Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nature Neuroscience*, pages 1–11, 2020.

[9] Robert M French. Catastrophic forgetting in connectionist networks. *Trends in cognitive sciences*, 3(4):128–135, 1999.

[10] Peter Jedlicka, Matus Tomko, Anthony Robins, and Wickliffe C Abraham. Contributions by metaplasticity to solving the catastrophic forgetting problem. *Trends in Neurosciences*, 2022.

[11] Larry Abbott and Karel Svoboda. Brain-wide interactions between neural circuits. *Current Opinion in Neurobiology*, 65(C), 2020.

[12] Herbert Jaeger and Harald Haas. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *science*, 304(5667):78–80, 2004.

[13] Wolfgang Maass, Prashant Joshi, and Eduardo D Sontag. Computational aspects of feedback in neural circuits. *PLoS Comput Biol*, 3(1):e165, 2007.

[14] David Sussillo and Larry F Abbott. Generating coherent patterns of activity from chaotic neural networks. *Neuron*, 63(4):544–557, 2009.

[15] Laureline Logiaco, L F Abbott, and Sean Escola. Thalamic control of cortical dynamics in a model of flexible motor sequencing. *Cell Reports*, 35(9):109090, 2021.

[16] Ta-Chu Kao, Mahdieh S Sadabadi, and Guillaume Hennequin. Optimal anticipatory control as a theory of motor preparation: A thalamo-cortical circuit model. *Neuron*, 109(9):1567–1581, 2021.

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

[18] Frank A Middleton and Peter L Strick. Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain research reviews*, 31(2-3):236–250, 2000.

[19] Narender Ramnani. The primate cortico-cerebellar system: anatomy and function. *Nature reviews neuroscience*, 7(7):511–522, 2006.

[20] Erik S Carlson, Avery Hunker, Stefan G Sandberg, Timothy M Locke, Julianne Geller, Abigail Schindler, Steven Thomas, Martin Darvas, Paul E M Phillips, and Larry S Zweifel. Catecholaminergic Innervation of the Lateral Nucleus of the Cerebellum Modulates Cognitive Behaviors. *Journal of Neuroscience*, 2021. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.2406-20.2021. URL https://www.jneurosci.org/content/early/2021/02/02/JNEUROSCI.2406-20.2021.

[21] Maedbh King, Carlos R Hernandez-Castillo, Russell A Poldrack, Richard B Ivry, and Jörn Diedrichsen. Functional boundaries in the human cerebellum revealed by a multi-domain task battery. *Nature neuroscience*, 22:1371–1378, 2019.

[22] James A Brissenden, Sean M Tobyne, Mark A Halko, and David C Somers. Stimulus-specific visual working memory representations in human cerebellar lobule VIIb/VIIIa. *Journal of Neuroscience*, 41(5):1033–1045, 2021.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[23] Zhenyu Gao, Courtney Davis, Alyse M Thomas, Michael N Economo, Amada M Abrego, Karel Svoboda, Chris I De Zeeuw, and Nuo Li. A cortico-cerebellar loop for motor planning. *Nature*, 563(7729):113, 2018.

[24] Ben Deverett, Mikhail Kislin, David W Tank, and Samuel S-H Wang. Cerebellar disruption impairs working memory during evidence accumulation. *Nature communications*, 10(1):1–7, 2019.

[25] Francois P Chabrol, Antonin Blot, and Thomas D Mrsic-Flogel. Cerebellar contribution to preparatory activity in motor neocortex. *Neuron*, 2019.

[26] Nuo Li and Thomas D Mrsic-Flogel. Cortico-cerebellar interactions during goal-directed behavior. *Current opinion in neurobiology*, 65:27–37, 2020.

[27] Hirokazu Tanaka, Takahiro Ishikawa, Jongho Lee, and Shinji Kakei. The Cerebro-Cerebellum as a Locus of Forward Model: A Review. *Frontiers in Systems Neuroscience*, 14:19, 2020. ISSN 1662-5137. doi: 10.3389/fnsys.2020.00019. URL https://www.frontiersin.org/article/10.3389/fnsys.2020.00019.

[28] Joseph Pemberton, Ellen Boven, Richard Apps, and Rui Ponte Costa. Cortico-cerebellar networks as decoupling neural interfaces. *Advances in Neural Information Processing Systems*, 34, 2021.

[29] Ellen Boven, Joseph Pemberton, Paul Chadderton, Richard Apps, and Rui Ponte Costa. Cerebro-cerebellar networks facilitate learning through feedback decoupling. *bioRxiv*, 2022.

[30] Herbert Jaeger. The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148(34):13, 2001.

[31] David Marr. From Trinity College, Cambridge. *J. Physiol*, 202:437–470, 1969.

[32] Christian Rössert, Sergio Solinas, Egidio D'Angelo, Paul Dean, and John Porrill. Model cerebellar granule cells can faithfully transmit modulated firing rate signals. *Frontiers in cellular neuroscience*, 8:304, 2014.

[33] Guillaume Bellec, Franz Scherr, Anand Subramoney, Elias Hajek, Darjan Salaj, Robert Legenstein, and Wolfgang Maass. A solution to the learning dilemma for recurrent networks of spiking neurons. *Nature communications*, 11(1):1–15, 2020.

[34] T D Aumann. Cerebello-thalamic synapses and motor adaptation. *The Cerebellum*, 1(1):69–77, 2002.

[35] Nicholas J Audette, Sarah M Bernhard, Ajit Ray, Luke T Stewart, and Alison L Barth. Rapid plasticity of higher-order thalamocortical inputs during sensory learning. *Neuron*, 103(2):277–291, 2019.

[36] Aparna Suvrathan, Hannah L Payne, and Jennifer L Raymond. Timing rules for synaptic plasticity matched to behavioral function. *Neuron*, 92(5):959–967, 2016.

[37] Jerome N Sanes, Bozhidar Dimitrov, and Mark Hallett. Motor learning in patients with cerebellar dysfunction. *Brain*, 113(1): 103–120, 1990.

[38] Kenji Yamamoto, Mitsuo Kawato, Shinya Kotosaka, and Shigeru Kitazawa. Encoding of movement dynamics by Purkinje cell simple spike activity during fast arm movements under resistive and assistive force fields. *Journal of neurophysiology*, 97(2): 1588–1599, 2007.

[39] Timothy J Ebner and Siavash Pasalar. Cerebellum Predicts the Future Motor State. *Cerebellum (London, England)*, 7(4):583, 2008.

[40] Frederic Lanore, N Alex Cayco-Gajic, Harsha Gurnani, Diccon Coyle, and R Angus Silver. Cerebellar granule cell axons support high-dimensional representations. *Nature Neuroscience*, 24(8):1142–1150, 2021.

[41] Jason M Christie and Michael A Gaffield. The cerebellum encodes and influences the initiation and termination of discontinuous movements. *bioRxiv*, 2021. doi: 10.1101/2021.06.24.449622. URL https://www.biorxiv.org/content/early/2021/06/24/2021.06.24.449622.

[42] Joshua Dacre, Matt Colligan, Thomas Clarke, Julian J Ammer, Julia Schiemann, Victor Chamosa-Pino, Federico Claudi, J Alex Harston, Constantinos Eleftheriou, Janelle M P Pakan, and Others. A cerebellar-thalamocortical pathway drives behavioral context-dependent movement initiation. *Neuron*, 2021.

[43] Alfonso Fasano, Simon E Laganiere, Susy Lam, and Michael D Fox. Lesions causing freezing of gait localize to a cerebellar functional network. *Annals of neurology*, 81(1):129–141, 2017.

[44] Jonathan F Marsden. Cerebellar ataxia. *Handbook of clinical neurology*, 159:261–281, 2018.

[45] Matthew G Perich, Juan A Gallego, and Lee E Miller. A neural population mechanism for rapid learning. *Neuron*, 100(4):964–976, 2018.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

[46] M S Milak, V Bracha, and J R Bloedel. Context-dependent modulation of cerebellar nuclear neurons related to the performance of specific movement segments. In *Soc Neurosci Abstr*, volume 20, page 1746, 1994.

[47] Hideaki Ogasawara, Tomokazu Doi, Kenji Doya, and Mitsuo Kawato. Nitric oxide regulates input specificity of long-term depression and context dependence of cerebellar learning. *PLoS Computational Biology*, 3(1):e179, 2007.

[48] Barbara Feulner, Matthew G Perich, Raeed H Chowdhury, Lee E Miller, Juan A Gallego, and Claudia Clopath. Small, correlated changes in synaptic connectivity may facilitate rapid motor learning. *Nature communications*, 13(1):1–14, 2022.

[49] Jennifer L Raymond and Javier F Medina. Computational principles of supervised learning in the cerebellum. *Annual review of neuroscience*, 41:233–253, 2018.

[50] Masao Ito. Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9(4):304, 2008.

[51] Ta-Chu Kao and Guillaume Hennequin. Neuroscience out of control: control-theoretic perspectives on neural circuit dynamics. *Current opinion in neurobiology*, 58:122–129, 2019.

[52] N Alex Cayco-Gajic and R Angus Silver. Re-evaluating circuit mechanisms underlying pattern separation. *Neuron*, 101(4):584–602, 2019.

[53] Baktash Babadi and Haim Sompolinsky. Sparseness and expansion in sensory representations. *Neuron*, 83(5):1213–1226, 2014.

[54] Suzana Herculano-Houzel. Coordinated scaling of cortical and cerebellar numbers of neurons. *Frontiers in neuroanatomy*, page 12, 2010.

[55] Timothy M. Locke, Marta E. Soden, Samara M. Miller, Avery Hunker, Cerise Knakal, Julia A. Licholai, Karn S. Dhillon, C. Dirk Keene, Larry S. Zweifel, and Erik S. Carlson. Dopamine D1 Receptor–Positive Neurons in the Lateral Nucleus of the Cerebellum Contribute to Cognitive Behavior. *Biological Psychiatry*, 84(6):401–412, 2018. ISSN 18732402. doi: 10.1016/j.biopsych.2018.01.019. URL https://doi.org/10.1016/j.biopsych.2018.01.019.

[56] Jens G Klinzing, Niels Niethard, and Jan Born. Mechanisms of systems memory consolidation during sleep. *Nature neuroscience*, 22(10):1598–1610, 2019.

[57] Richard F Lewis. Context-dependent adaptation of visually-guided arm movements and vestibular eye movements: role of the cerebellum. *The Cerebellum*, 2(2):123–130, 2003.

[58] James B Heald, Máté Lengyel, and Daniel M Wolpert. Contextual inference underlies the learning of sensorimotor repertoires. *Nature*, 600(7889):489–493, 2021.

[59] Dimitar Kostadinov and Michael Häusser. Reward signals in the cerebellum: origins, targets, and functional implications. *Neuron*, 2022.

[60] Benjamin W Corrigan, Roberto A Gulli, Guillaume Doucet, Megan Roussy, Rogelio Luna, Kartik S Pradeepan, Adam J Sachs, and Julio C Martinez-Trujillo. Distinct neural codes in primate hippocampus and lateral prefrontal cortex during associative learning in virtual environments. *Neuron*, 2022.

[61] Paul G Anastasiades, David P Collins, and Adam G Carter. Mediodorsal and ventromedial thalamus engage distinct L1 circuits in the prefrontal cortex. *Neuron*, 109(2):314–330, 2021.

[62] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[63] James S Albus. A theory of cerebellar function. *Mathematical Biosciences*, 10(1):25–61, 1971. ISSN 0025-5564. doi: https://doi.org/10.1016/0025-5564(71)90051-4. URL http://www.sciencedirect.com/science/article/pii/0025556471900514.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Methods

Model architecture and training

	No feedback	Readout feedback	Cerebellar feedback	No feedback (cerebellar readout)
$\mathbf{h}_t$	$\alpha\mathbf{h}_{t-1} + W_{hh}f(\mathbf{h}_{t-1}) + W_{ih}\mathbf{x}_t$	$\alpha\mathbf{h}_{t-1} + W_{hh}f(\mathbf{h}_{t-1}) + W_{ih}\mathbf{x}_t + W_{zh}\mathbf{z}_t$	$\alpha\mathbf{h}_{t-1} + W_{hh}f(\mathbf{h}_{t-1}) + W_{ih}\mathbf{x}_t + W_{\mathcal{C}h}\mathbf{c}_t$	$\alpha\mathbf{h}_{t-1} + W_{hh}f(\mathbf{h}_{t-1}) + W_{ih}\mathbf{x}_t$
$\mathbf{z}_t$	$W_{\text{rdt}}f(\mathbf{h}_t)$	$W_{\text{rdt}}f(\mathbf{h}_t)$	$W_{\text{rdt}}f(\mathbf{h}_t)$	$\mathcal{C}(f(\mathbf{h}_t))$
$\mathbf{c}_t$	NA	NA	$\mathcal{C}(f(\mathbf{h}_{t-1}))$	NA

Table S1. Dynamics of the different model variants, where $\mathbf{h}_t$ is the cortical RNN state, $\mathbf{z}_t$ the readout and $\mathbf{c}_t$ cerebellar feedback. For the experiments presented here we set $f = \tanh$ and $\mathcal{C}$ is the cerebellar feedforward network with one hidden layer, $\mathcal{C}(f(\mathbf{h})) = W_{\text{PF}}f^{\mathcal{C}}(W_{\text{MF}}f(\mathbf{h}))$. W_{hh}, RNN recurrent weights; W_{ih}, stimulus-to-RNN weights; W_{rdt}, (cortical) readout weights; $W_{\mathcal{C}h}$, cerebellar-to-RNN weights; W_{MF}, cerebellar mossy fibre weights; W_{PF}, cerebellar parallel fibre weights; $f^{\mathcal{C}}$ set as ReLU.

The complete dynamics of each model architecture that we consider (Extended Data Fig. S1; no feedback, readout feedback, cerebellar feedback, no feedback with cerebellar readout) are given in Table S1. In all of our simulations we use a recurrent neural network (RNN) with 50 time-discrete units (see section below).

Unless otherwise stated, the feedforward cerebellar network contains a single hidden layer with 1000 units (granule cells), but other hidden layer sizes are also considered (Figs. 2d and 4e). This yields a divergence from the cortical RNN to the cerebellar granular layer of 50:1000 = 1:20. The cerebellar output layer, which we interpret as Purkinje cells, on the other hand, mirrors the desired task outcome and is therefore of significantly lower dimensionality (3 in evidence accumulation task and 2 in all other tasks).

For each task simulation, network parameters are initialised as follows. The RNN input, recurrent and cerebellar feedback weights W_{ih}, W_{hh}, $W_{\mathcal{C}h}$ are drawn from a uniform distribution $W^{\text{init}} \sim \mathcal{U}(-a, a)$ where $a = \frac{1}{\sqrt{\text{size(RNN)}}} = \frac{1}{\sqrt{50}}$. The readout weights W_{rdt} and cerebellar weights, W_{MF}, W_{PF}, are initialised as $W^{\text{init}} \sim \mathcal{U}(-a, a)$, where a is normalised according to the respective number of input neurons, $a = \frac{1}{\sqrt{\text{size(input)}}}$. Note that the readout weight W_{rdt} includes a bias term. We conducted each task simulation with 5 random seeds for initialisation, except for the theoretical analysis done in Fig. 7b,c and Extended Data Fig. S17 for which we use 10.

During the learning of a task model parameters are updated using gradient descent from the task error signal $E = \sum_t E_t$ with respect to to the model parameters (see section below). For each dataset each training session covers 1000 random examples, presented to the model in batch sizes of 10 which we call a "trial". The test set (used after training) also covers 1000 randomly generated examples. When analysing the learned network dynamics (e.g. model output with and without cerebellar ablation) the model with the best validation error during training was selected. An ADAM optimiser[62] was used with initial learning rate $\eta = 0.001$ for the RNN (when plastic), readout and cerebellar network, except for the delayed association task for which we found an RNN learning rate of $\eta = 0.0025$ to provide more stable learning. The different plasticity constraints of the entire model - termed "fixed RNN", "input plastic", and "fully plastic" - are defined with respect to the cortical parameters of Eq. 1 as follows. For the fixed RNN case, only the cortical readout weights W_{rdt} are learned. For the input plastic case, RNN input weights and W_{ih} and $W_{\mathcal{C}h}$ are also learned. Finally, for the fully plastic case, the recurrent weight W_{hh} is also learned. In all of these cases the cerebellar "parallel fibres" W_{PF} are learned, whilst the "mossy fibres" W_{MF} remain constant, in line with mossy fibres synapses being (relatively) stable[31,63].

Continuous dynamics of RNN model

A continuous version of our RNN can be expressed as

$$\begin{aligned} \tau_M \dot{\mathbf{h}} &= -\mathbf{h} + R_m\left(W_{hh}f(\mathbf{h}) + W_{ih}\mathbf{x} + W_{\mathcal{C}h}\mathbf{c}\right) \\ \mathbf{z} &= W_{\text{rdt}}f(\mathbf{h}) \end{aligned} \tag{3}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

where τ_M is the membrane time constant (not to be confused with the cerebellar time window τ), R_m is the membrane resistance, and f is the rate-based non-linearity which we set as $f = \tanh$. Discretising Eqs. 3 with timesteps of Δt yields equations in Table S1, where $\alpha = \exp(-\frac{\Delta t}{\tau_M})$. Note that as in[33] we ignore the $(1-\alpha)R_m$. This simplifies notation and has no effect on dynamics if model weights are scaled accordingly. In general we use $\tau_M \approx 20$ms and $\Delta t = 50$ms for the drawing tasks (Figs. 2,3 and 4) and a higher $\tau_M \approx 90$ms with $\Delta t = 200$ms for the cognitive tasks (Figs. 5,6,7 and 8) in line with[6]). In both cases this gives us a cortical internal memory $\alpha = 0.1$.

Cortical and cerebellar learning rules

When the desired task outcome $\mathbf{y}_t$ is provided the associated error is computed as $E_t = \mathcal{E}(\mathbf{z}_t, \mathbf{y}_t)$ for the cortical network and $E_t^{\mathcal{C}} = \mathcal{E}(\mathbf{c}_{t-\tau}, \mathbf{y}_t)$ for the cerebellar network, where $\mathcal{E}$ denotes the task error function (mean squared error and cross-entropy loss for regression and classification tasks respectively) and τ is the cerebellar time window. The error gradients for the readout and cerebellar weights W_{rdt}, W_{PF} can then be obtained locally with a simple delta-rule on the gradient of the error signal. That is,

$$\begin{aligned} \Delta W_{\text{rdt}} &= \eta \frac{dE}{dW_{\text{rdt}}} = \eta \frac{dE_t}{d\mathbf{z}_t} f(\mathbf{h}_t)^\top \\ \Delta W_{\text{PF}} &= \eta \frac{dE_t^{\mathcal{C}}}{dW_{\text{PF}}} = \eta \frac{dE^{\mathcal{C}}}{d\mathbf{c}_{t-\tau}} \mathbf{GC}_{t-\tau}^\top \end{aligned} \tag{4}$$

Where η denotes the learning rate of the cortico-cerebellar network and **GC** denotes the hidden granule cell activity of the cerebellar network which is computed as $\mathbf{GC}_t = f^{\mathcal{C}}\left(W_{\text{MF}} f(\mathbf{h}_{t-1})\right)$ (cf. Eq. 2).

For the input/recurrent weights W_{ih}, $W_{\mathcal{C}h}$, W_{hh} - when plastic - obtaining error gradients is more difficult as temporal dependencies need to be considered. To improve biological feasibility in this work we avoid backpropagation through time (BPTT) and instead use the eprop algorithm[33]. Details can be found in[33], but the main idea is that BPTT can be approximated with a mixture of locally computed synaptic eligibility traces and current learning signal. Specifically, the error gradient for a given synapse w_{ji} from neuron i to j is computed as

$$\Delta w_{ji} = \eta \frac{dE}{dw_{ji}} = \eta \sum_t L_j^t \epsilon_{ji}^t \tag{5}$$

where for ease of notation we now use the superscript to denote timestep t and $L_j^t = \frac{dE_t}{dh_j^t}$ is the neuron j learning signal (obtained by one-step backpropagation through space except for the cerebellar readout architecture in Extended Data Fig. S1D). ϵ_{ji}^t is the synaptic eligibility trace of w_{ji} which is defined recursively by

$$\epsilon_{ji}^t = \frac{\partial h_j^t}{\partial h_j^{t-1}} \epsilon_{ji}^{t-1} + \frac{\partial h_t}{\partial w_{ji}} \tag{6}$$

where ϵ_{ji}^0 is initialised as zero. Note that the terms in Eq. 6 are locally available to the synapse. In the case of our network dynamics (Eq. 1), the eligibility trace is simply defined by $\epsilon_{ji}^t = \alpha \epsilon_{ji}^{t-1} + a_i$, where a_i is the activation of the presynaptic neuron i (e.g. $\tanh(h_i)$ or c_i).

For all weights, the error gradients are accumulated across multiple examples (i.e. batch update) and timesteps before the weights themselves are updated.

Learning rules for cerebellar-to-cortical consolidation

A period of "consolidation" is considered for the trained models of the delayed association task (Fig. 8 and Extended Data Fig. S18). During this period the model is presented with further trials (batch size 10) of training data but without their associated targets. The forward dynamics of the model then run as normal (Eq. 1) but now we use a consolidation learning rule for the RNN weights. We consider both an optimal learning rule which uses the least-squares algorithm and also a simple biological learning rule.

We first present the optimal consolidation learning rule, since this motivates the biological rule. We want to change the recurrent (cortico-cortical) input to match the cerebellar-cortico input over the task. To this end we concatenate the time-dependent RNN activities $\mathbf{H} = \bigoplus_{t \geq 1} \mathbf{h}_t$ and cerebellar output activities $\mathbf{C} = \bigoplus_{t \geq 1} \mathbf{h}_t$, where $\oplus$ denotes vector

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

concatenation. We then set the change in recurrent weight $\Delta^{\text{cons}} W_{hh}$ with $\Delta^{\text{cons}} W_{hh} = \eta^{\text{cons}}_{\text{RNN}} F^{\text{lsq}}$ where $\eta^{\text{cons}}_{\text{RNN}}$ is the RNN consolidation learning rate and F^{lsq} is the least-squares solution

$$F^{\text{lsq}} f(\mathbf{H}) = W_{\mathcal{C}h}\mathbf{C} \tag{7}$$

At the same time the cerebellar-cortical weights $W_{\mathcal{C}h}$ decay according to

$$\Delta W^{\text{cons}}_{\mathcal{C}h} = -\eta^{\text{cons}}_{\mathcal{C}h} W_{\mathcal{C}h} \tag{8}$$

Where $\eta^{\text{cons}}_{\mathcal{C}h}$ is the rate of cerebellar-cortical decay. In the experiments shown we select $\eta^{\text{cons}}_{\mathcal{C}h} = \eta^{\text{cons}}_{\text{RNN}} = 0.1$.

For the biological learning rule, the cerebellar-cortical weight decays as in Eq. 8 but now the RNN weights are updated according to the ratio of cerebellar feedback against the whole population activity. That is, for the recurrent weight from neuron i to neuron j we have

$$\Delta^{\text{cons}} w_{ij} = \eta^{\text{cons}}_{\text{RNN}} \frac{\text{cerebellar input to } j}{\text{total RNN activity}} = \eta^{\text{cons}}_{\text{RNN}} \frac{W^j_{\mathcal{C}h}\mathbf{c}_t}{\sum_k f(h_{k,t})} \tag{9}$$

For arbitrary timestep t and where $W^j_{\mathcal{C}h}$ denotes the jth row of the cerebello-cortical weight $W_{\mathcal{C}h}$.

To demonstrate that Eq. 9 leads to changes in cortico-cortico input which are proportional to the cerebello-cortical input, we see that the change in recurrent input to a given RNN neuron j at time t becomes

$$\begin{aligned}
\underbrace{\Delta W^j_{hh} f(\mathbf{h}_t)}_{\Delta \text{ recurrent input}} &= \sum_{i \in \text{RNN}} \Delta w^{(hh)}_{ij} f(h_{i,t}) \\
&= \sum_{i \in \text{RNN}} \eta^{\text{cons}}_{\text{RNN}} \left(\frac{W^j_{\mathcal{C}h}\mathbf{c}_t}{\sum_{k \in \text{RNN}} f(h_{k,t})} \right) f(h_{i,t}) \\
&\propto \frac{1}{\sum_{k \in \text{RNN}} f(h_{k,t})} W^j_{\mathcal{C}h}\mathbf{c}_t \sum_{i \in \text{RNN}} f(h_{i,t}) \\
&= \underbrace{W^j_{\mathcal{C}h}\mathbf{c}_t}_{\text{cerebellar input}}
\end{aligned}$$

That is, we recover a solution (up to proportionality) to Eq. 7. For this biological learning rule, to improve network stability, we found it beneficial to increase the RNN consolidation learning rate such that $\eta^{\text{cons}}_{\text{RNN}} = 3\eta^{\text{cons}}_{\mathcal{C}h} = 0.3$ (where $\Delta^{\text{cons}} w_{ij}$ is accumulated over the whole sequence). This explains the initially faster learning (over the first few trials) for the biological learning rule (Fig. 8e).

For this consolidation learning period a learning optimiser is not used (i.e. ADAM is not used). Note that these consolidation learning rules do not require information about the desired task outcome (i.e. target) and are in that sense unsupervised.

Task details

1. Line drawing task

For the line drawing task, the model has to transform one of six possible 10-dimensional binary inputs $\mathbf{x} \in [0,1]^{10}$ at timestep 1 into an associative "go" 2-dimensional line $\mathbf{y}^{\text{line}}$ (for five of the inputs) or a "no-go" stay at the origin (for one of the inputs). The starting point for each line is the origin, and the endpoints of each line are evenly spaced on the edge of the unit circle (see Fig. 2a, black dashed line). The model learns to draw the line over 20 discrete timesteps, with the intermediate target points spaced evenly, i.e. for a line with endpoint y_{end} we have $\mathbf{y}^{\text{line}} = \{0, y_1, y_2, \dots, y_{\text{end}}\} = \{0, \frac{1 \times y_{\text{end}}}{19}, \frac{2 \times y_{\text{end}}}{19}, \dots, \frac{19 \times y_{\text{end}}}{19}\}$.

For the stimulus timestep (timestep 1) as well as the remaining 19 timesteps, the model receives (through its W_{ih} connection) zero-mean Gaussian noise $\xi \sim \mathcal{N}(0; 0.1^2)$. Model errors are computed as the mean-squared error to the target response. Unless otherwise stated a cerebellar time window $\tau = 3$ timesteps ($\approx$ 150ms when $\alpha = 0.1$) is used. The prediction error across time delay t_0 between cortical output and cerebellar (or cortical) output (Fig. 2e) is computed as the cue/time average $||\mathbf{c}_{t+t_0} - \mathbf{z}_t||$, where $||.||$ is the Euclidean norm.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

To analyse the effects of cerebellar ablation we consider partial cerebellar ablation at the start, middle, and end of the sequence (Fig. 2h-k and Extended Data Fig. S5). The specific time windows of these ablation periods are timesteps [1-6, 8-13, 15-20] (inclusive), respectively.

Curl-field variant: Once the models of the line drawing task are trained, we tested whether they could re-translate the same external inputs to a curl-field variant of the task (see e.g.[45]). For this we selected models with cortical internal memory $\alpha = 0.5$, since we found this resulted in faster learning which was comparable to the presented experimental data[45], but we find $\alpha = 0.1$ (as presented in Fig. 2) also learns but more slowly. Switching and learning this curl-field new task "context" involved retraining the models to new desired outcomes (central grey curves in Fig. 3c).

Specifically, the curl-field target responses have the same end-point for each line (or same "no-go" zero cue), but intermediate target points now form a semi-ellipse between the origin and the respective end-point. Given the desired endpoint $\mathbf{y}_{\text{end}} = \begin{pmatrix} y^0_{\text{end}} \\ y^1_{\text{end}} \end{pmatrix}$, this can be parameterised by

$$\mathbf{y}_t = \begin{pmatrix} y^0_t \\ y^1_t \end{pmatrix} = \begin{pmatrix} \frac{y^0_{\text{end}}}{2} + \frac{1}{2}\cos\theta\cos t - \frac{1}{2}\sin\theta\sin t \\ \frac{y^1_{\text{end}}}{2} + \frac{1}{2}\sin\theta\cos t + \frac{1}{2}\cos\theta\sin t \end{pmatrix} \tag{10}$$

where $\theta = \arctan(\frac{y^1_{\text{end}}}{y^0_{\text{end}}})$ is the angle to the end point and t runs uniformly between 0 and π (or, for direction towards $(x_{\text{end}}, y_{\text{end}})$ as in our experiments, from π to 2π).

To test how context-dependent cerebellar processing could enable rapid task switching, we considered the extent to which parallel fibre (PF) weights are shared across task contexts. In particular, we label the percentage of PFs used for each context as the PF task overlap. For example, if the PF task overlap is 25%, then 25% of the PFs used for cerebellar processing apply to both task contexts, whilst 75% specifically apply (and are trained) to the current context. Before learning, the PFs which are not shared (i.e. only apply to the curl-field context) are initialised randomly as in the original line-drawing task.

Neuronal activity and covariance during task switching: The change in activities and change in covariances (Fig. 3d-f and Extended Data Fig. S7) are computed as in[48]. We record the RNN time-dependent activities (post non-linearity) given 1000 input examples in multiple periods: task 1 baseline, task 2 and task 1 switching (Fig. 3a). For the latter two periods these are recorded at their respective end, whilst we take two samples of the baseline period at its start and end. The change in activity between any two periods $P1$ and $P2$ is the average change in activity for a given neuron i, which is given by

$$\Delta^{(P1,P2)}_{\text{act}} h_i = \frac{|h^{P2}_i - h^{P1}_i|}{\text{std}_i} \tag{11}$$

where h^{P1}_i, h^{P2}_i are the time-varying input-dependent activities of neuron i for periods $P1$, $P2$ respectively, and std_i is the standard deviation of that neuron in the start of the task 1 baseline period. Here $|.|$ denotes the average (absolute) difference in activity across timesteps and input examples.

For each period, we also compute the covariance matrix of the RNN population. The change in covariance between two sessions $\Delta^{(P1,P2)}_{\text{cov}}$ is then computed as 1 minus the Pearson correlation between their respective covariance matrices[48].

For the task 2 and task 1 switching periods we report changes with respect to the start of the task 1 baseline period. To account for natural variability in the network and better compare to the neural data in[45], we normalise the changes by taking away the changes observed within the baseline period itself. For example, the change in covariance in the task 2 period is $\Delta^{(B1,T2)}_{\text{cov}} - \Delta^{(B1,B2)}_{\text{cov}}$, where $B1$, $B2$, $T2$ are the start of the task 1 baseline, end of task 1 baseline, and (end of) task 2 respectively. We apply the same normalisation to the reported experimental changes in the monkey M1 and PmD[45]; this normalisation leads to (average) near-zero change for the M1 activity and PmD (Fig. 3f).

2. Digit drawing task

For the digit drawing task the inputs are the same as the 10-dimensional binary vectors used in the line drawing task, except now the model must draw an associative digit over 20 timesteps instead of line (Fig. 4a). The targets $\mathbf{y}^{\text{digit}}$ are constructed manually within the space $[0, 1]^2$ and resemble the digits from 0 to 5 (inclusive). For exact implementation refer to the provided code (see below).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

For the standard model with cerebellar feedback a cerebellar time window $\tau = 3$ timesteps ($\approx$ 150ms when $\alpha = 0.1$) is generally used. For the model using cerebellar feedback with a temporal basis, we model the cerebellum with a range of time windows, i.e. $\tau = \{\tau_i\}_i$ for some distinct $\tau_i \geq$ 0ms. In this task we consider $\tau = \{\tau_i\}_{i=0}^{i=5}$ with $\tau_i = i$ timesteps (i.e. 0-250ms), so that the final cerebellar output is a concatenation of task predictions which span over the proceeding 250ms period. Explicitly, after training we have cerebellar feedback, $\mathbf{c}_t \approx \bigoplus_{i=0}^{i=5} \mathbf{y}_{t+i}$, where $\oplus$ denotes vector concatenation.

Zero-mean Gaussian noise $\xi \sim \mathcal{N}(0; 0.1^2)$ is added to the input at each timestep. Model errors are computed as the mean-squared error to the target response.

To analyse the effects of cerebellar ablation we consider the same partial cerebellar ablation periods as in the line-drawing task. That is, we consider cerebellar ablation at the start, middle, and end of the sequence (Fig. 4 and Extended Data Fig. S9), which correspond to timesteps [1-6, 8-13, 15-20] (inclusive), respectively.

3. Evidence accumulation task

In the evidence accumulation task the model receives 2-dimensional binary inputs (i.e. $\mathbf{x} \in [0, 1]^2$) over a presentation period of $T^{\text{pres}} = 45$ timesteps. A non-zero input can occur for at most one of the two dimensions; that is, $x_t \in \{(1 \quad 0)^\top, (0 \quad 1)^\top, (0 \quad 0)^\top\}$, where the rate of zero inputs $x_t = (0 \quad 0)^\top$ defines the sparsity of input ρ ($\rho = 0.7$ in our simulations). After this presentation of input there is then a delay period of $T^{\text{del}} = 5$ timesteps after which the model must classify at which dimension more non-zero input was received (or whether the number at each dimension was the same). That is, the desired outcome y takes one of three values which respectively correspond to more input in the first dimension, more input in the second dimension, or the same. This task resembles the experimental structure of[24], in which mice were trained to select the side of their whiskers which received more air puffs .

Zero-mean Gaussian noise $\xi \sim \mathcal{N}(0; 0.1^2)$ is added to the input at each timestep. Model errors are defined by the cross-entropy loss to the target response.. Model "belief" (Figs 5D and S11) is defined as the model probability (obtained by applying a softmax on the readout) of the correct classification. Unless otherwise stated a cerebellar time window $\tau = 3$ timesteps ($\approx$ 600ms when $\alpha = 0.1$) is used. For both readout and cerebellar feedback models, we apply a softmax operation to the feedback returned to the RNN so as to bound its values between 0 and 1.

To analyse the effects of cerebellar ablation we consider full cerebellar ablation (for the entire sequence 1-50; see Fig. 5d and Extended Data Fig. S11a-c, left) and also partial periods of ablation: at the start, middle, and end of the sequence (Fig. 5e,f and Extended Data Fig. S11a-c, right). The specific time windows of these partial ablation periods are timesteps [1-15, 15-30, 30-45] (inclusive), respectively. To improve readability of our results, the mean error presented in the training curves for this task is smoothed using a Savitzky-Golay filter with window length 25 and polynomial order 3.

To compute the dependence of model choice on inputs over different temporal bins (Fig. 5f), we follow the method in[24]. In particular, we divide the presentation period evenly into 3 time windows - [1-15, 16-30, 31-45] - and fit the model choice according to a logistic regression model

$$\hat{y} = S(\beta_1 E_1 + \beta_2 E_2 + \beta_3 E_3) \tag{12}$$

Where $\hat{y}$ denotes the predicted model choice probability, S is the sigmoid logistic function, $E_i = \#R_i - \#L_i$ is the different in the total number of 'right' and 'left' inputs in window i, and β_i is the respective weight on that window. $\hat{y}$ is fitted to minimise the negative log likelihood of the observed model decisions. We present the normalised weights of each window $\frac{\beta_i}{\beta_1+\beta_2+\beta_3}$.

History-centric cases: In line with[24], we observed cerebellar ablation to be particularly detrimental to input examples for which correct classification would depend on adequately maintaining past inputs (Fig. 5e,f and Extended Data Fig. S11), which we refer to as "history-centric" examples. We defined an input example as being history-centric if exposure only to the final third of the input sequence would lead strictly to the wrong answer. That is, examples $(\mathbf{x}, y)$ such that the "final-third target" $y_{\text{final third}} = \max(\sum_{t=33}^{T=50} x_t^0, \sum_{t=33}^{T=50} x_t^1)$ is not equal to the desired outcome $y = \max(\sum_{t=1}^{T=50} x_t^0, \sum_{t=1}^{T=50} x_t^1)$.

4. Delayed association task

In the delayed association task the model must associate one of two 10-dimensional binary inputs at timestep 1 to a desired binary response y at timestep T, where T is the sequence length or "delay" period[23]. We select $T = 15$

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

timesteps but also consider other lengths (Fig. 7a). The task error (as presented in the main text) is defined at the end of the sequence. For stability, we train the network output 5 timesteps from the end of the sequence (timestep 10 onwards when $T = 15$).

Zero-mean Gaussian noise $\xi \sim \mathcal{N}(0; 0.1^2)$ is added to the input at each timestep. Model errors are defined by the cross-entropy loss to the target response. Model "selectivity" is defined as the model output (readout) at the dimension of the correct classification (prior to the softmax operation). Unless otherwise stated a cerebellar time window $\tau = 3$ timesteps ($\approx$ 600ms when $\alpha = 0.1$) is used. For both readout and cerebellar feedback models, we apply a softmax operation to the feedback returned to the RNN so as to bound its values between 0 and 1.

To analyse the effects of cerebellar ablation we consider cerebellar ablation within a particular time window between timesteps 8-12 (inclusive) which approximately mirrors the timings in [23] (Fig. 6d,e and Extended Data Fig. S16) and also partial ablation periods during the start, middle, and end of the sequence (Fig. 6f). The specific time windows of these partial ablation periods are timesteps [1-5, 6-10, 11-15], respectively. To improve readability of our results, the mean error presented in the training curves for this task is smoothed using a Savitzky-Golay filter with window length 25 and polynomial order 3.

Control-theoretic estimation of cerebellar feedback

For the delayed association task we analyse cerebellar-to-cortical input from a control-theoretic point of view. In particular, we quantify the effect of plasticity in the pathway between the cerebellar network and cortical RNN ($W_{\mathcal{C}h}$) on cortical activations by estimating the energy cerebellar feedback induces in RNN state space [51]. This level of energy reflects the potency of feedback onto the RNN: a low energy would reveal a suppressed RNN response, whereas a high energy would reveal an amplified response. We speculated that these two cases would arise from a non-optimised $W_{\mathcal{C}\langle}$ and optimised $W_{\mathcal{C}\langle}$, respectively (Extended Data Fig. S13A).

As per Kao and Hennequin [51], we compute the energy of cerebellar feedback through the *controllability Gramian* P associated with RNN dynamics. Informally, P describes the "intrinsic manifold" of the RNN and describes the directions in which the RNN is most (or least) likely to visit. Formally, given a direction $\mathbf{v}$ in state space, the average energy generated along direction v is

$$\sigma(\mathbf{v}) = \mathbf{v}^\top P \mathbf{v} \tag{13}$$

In general, the Gramian matrix P is only defined for linear systems. In this work we therefore generalise the notion of controllability for the non-linear RNN dynamics as defined in Eq. 1. Here we use the noise covariance matrix Σ in its place, which for linear systems is shown to be equivalent to the Gramian, $\Sigma = P$ [51]. Explicitly, we compute Σ as the time-course average covariance of RNN hidden activations h_t under noisy inputs which follow a Wiener process. That is, $\Sigma = \mathbb{E}_t[\text{cov}(H_t)]$ where $H_t = \{\mathbf{h}_t^1, \mathbf{h}_t^2, \cdots, \mathbf{h}_t^N\}$ is a set of N samples of RNN states which each evolve according to

$$\mathbf{h}_t = \alpha \mathbf{h}_{t-1} + W_{hh} f(\mathbf{h}_{t-1}) + W_{\mathcal{C}h}\mathbf{c}_t + \xi_t; \quad \xi_{i,t} \sim \text{Wiener process} \tag{14}$$

In our experiments we use $N = 500$ samples and simulate Eq. 14. To ignore intrinsic RNN transients that occur at the start of simulation, we discard the RNN states during the first 5 simulation timesteps when computing Σ. The energy generated from cerebellar feedback is then $\sigma(\mathbf{h}^{\mathcal{C}}) = (\mathbf{h}^{\mathcal{C}})^\top \Sigma \mathbf{h}^{\mathcal{C}}$, where $\mathbf{h}^{\mathcal{C}} = \frac{W_{\mathcal{C}h}\mathbf{c}}{||W_{\mathcal{C}h}\mathbf{c}||}$ is the normalised direction being driven by the cerebellum in RNN state space. We report the energy generated (during the noise dynamics of Eq. 14) by cerebellar feedback at timestep 10, a time chosen strictly after the initial RNN transient phase (Extended Data Fig. S13b). For comparison we compare this to the energy generated by 100 random sample directions $\mathbf{v} \sim \mathcal{N}(0, I)$ where I is the identity matrix. To enable greater interpretability we then normalise these energies by its highest possible value $\max_{||\mathbf{v}||=1} \mathbf{v}^\top \Sigma \mathbf{v}$; i.e. the input which elicits maximal amplification of RNN dynamics. This value can be computed as $\mathbf{u}^\top \Sigma \mathbf{u}$ where $\mathbf{u}$ is the principal eigenvector of Σ.

Cerebellum as a decoder of fading memories

For the delayed association task we discussed the need for a greater number of hidden cerebellar units (granule cells) to achieve good task performance (Fig. 7a). In particular, we find that the number of granule cells (GCs) required is inversely proportional to the *signal-to-noise* (SNR) of the RNN hidden neurons.

$$\#\ \text{GCs needed} \propto \frac{1}{\sqrt{\text{SNR(RNN)}}} \tag{15}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Under reasonable assumptions, Eq. 15 can be shown analytically. We go through the main steps of this proof; a complete derivation can be found in supplementary note 1.

First, we note that we express the RNN activations $f(\mathbf{h})$ as the sum of "task-agnostic" and "task-relevant" features, or components. The former might incorporate general network connectivity, or intrinsic neural noise, whilst the latter depends solely on the task condition. We take these components to be independent of each other, so that the total variance of RNN activations $\sigma^2_{\text{rnn}} = \text{Var}(f(\mathbf{h}))$ can then be expressed as

$$\sigma^2_{\text{rnn}} = \sigma^2_\omega + \sigma^2_\zeta \tag{16}$$

Where σ^2_ω and σ^2_ζ denote the variance caused by the task-agnostic and task-relevant components, respectively. Assuming each component is centred around zero (which we observe using simulations), the ratio of these two quantities then represents the SNR of the RNN population, $\text{SNR(RNN)} = \frac{\sigma^2_\zeta}{\sigma^2_\omega}$. With a further assumption that this RNN population is centered at zero (which we observe as reasonable) and the mossy fibres W_{MF} are normally distributed, it can then be shown that the general cerebellar GC input I_{GC} can be expressed as

$$I_{\text{GC}} = I_\omega + I_\zeta \;\; I_\omega \sim \mathcal{N}(0,\, (k\sigma^2_{\text{MF}})\sigma^2_\omega),\;\; I_\zeta \sim \mathcal{N}(0,\, (k\sigma^2_{\text{MF}})\sigma^2_\zeta) \tag{17}$$

Where k is the number of neurons in the RNN population and σ^2_{MF} is the variance of the mossy fibre weights, $\sigma^2_{\text{MF}} = \text{Var}(W_{\text{MF}})$. Given this distribution for the input, the key step is considering the probability that the task-relevant input I_ζ can drive a sub-threshold GC activation ($I_{\text{GC}} < 0$) in one task-setting but super-threshold activation ($I_{\text{GC}} > 0$) in the other. We term this as the probability of a "unique activation"; it can then be shown that for small SNR values

$$P(\text{unique activation}) \approx \frac{\sqrt{2}}{\pi}\frac{(k\sigma^2_{\text{MF}})\sigma_\zeta}{(k\sigma^2_{\text{MF}})\sigma_\omega} = \frac{\sqrt{2}}{\pi}\frac{\sigma_\zeta}{\sigma_\omega} = \frac{\sqrt{2}}{\pi}\sqrt{\text{SNR(RNN)}} \tag{18}$$

In other words, whether or not there is a unique activation in the GC is simply governed by SNR-depedent chance. Intuitively, therefore, given a large enough GC layer the likelihood is that at least one GC will have a unique activation. In fact, it can be shown that the number of neurons needed to have at least one unique activation with confidence θ can be expressed as (which is the result cited in the main text)

$$\#\text{ GCs needed} = -\frac{\pi}{\sqrt{2}}\log(1-\theta)\frac{1}{\sqrt{\text{SNR(RNN)}}} \tag{19}$$

To estimate SNR(RNN) in the models for the delayed association task (Fig. 7b, left axis), we compute the variance of the task-agnostic component as the (average) variance of the population under the same task stimulus s, i.e. $\sigma^2_\omega = \mathbb{E}_s[\text{Var}(f(\mathbf{h})|s]$. Be equally calculating the total variance $\sigma^2_{\text{rnn}} = \text{Var}(f(\mathbf{h}))$, the variance of the task-relevant component is then simply computed as the difference to the total variance, i.e. $\sigma^2_\zeta = \sigma^2_{\text{rnn}} - \sigma^2_\omega$. To determine the minimum number of granule cells required to decode the stimulus from the RNN activity (Fig. 7b, right axis), we tested whether the cerebellar network could be trained to successfully discriminate the stimulus after 40 training sessions for varying quantities of granule cells (quantities as described below). The cerebellar network was deemed to successfully decode the stimulus if, for at least 9 of the 10 seeds, the average error during the last 4 training sessions was less than 5%.

We also generated synthetic inputs to directly test the ability of the cerebellar network for different SNRs (Fig. 7c and Extended Data Fig. S17). In this case the network must classify one of two 10-dimensional input of the form $\mathbf{x} = \mathbf{x}_\omega + \mathbf{x}_\zeta$. Both $\mathbf{x}_\omega$ and $\mathbf{x}_\zeta$ are drawn randomly from normal distributions of zero mean, but only $\mathbf{x}_\zeta$ changes for each task condition (i.e. once sampled $\mathbf{x}_\omega$ is fixed). The distributions $\mathbf{x}_\omega$, $\mathbf{x}_\zeta$ have distinct variances σ^2_ω, σ^2_ζ respectively. The SNR of the input data is then simply $\frac{\sigma^2_\zeta}{\sigma^2_\omega}$, where we set $\sigma^2_\omega = 1$ and vary σ^2_ζ across a range of values (see below). The architecture of the cerebellar network remains the same as in the main cortico-cerebellar model, except that a spiking activation function is used at the granular layer; that is, $f_{\mathcal{C}} = f_{\text{spike}}$ in Eq. 2 where $f_{\text{spike}}(I) = 1$ if $I > 0$ and 0 otherwise.

To verify whether Eq. 19 indeed holds empirically, we ran the network on the dataset described above over a range of different SNRs s and network sizes m. We considered $s = (\frac{1}{2})^k$ and $m = 2^k$ for k between 0 and 14 inclusive (i.e. 225 distinct SNR/size configurations). We ran each configuration over 10 random seeds. To determine the number of GCs needed for a given SNR s, we considered the minimum m needed such that 1. the GC population vector is distinct for the two task inputs $\mathbf{x}_1$, $\mathbf{x}_2$ (Fig. 7c and Extended Data Fig. S17a) and correspondingly 2. the cerebellar readout (Purkinje cells) can successfully learn to classify the initial input (Extended Data Fig. S17a). Classification was again

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

deemed successful if the average error during the last 4 training sessions out of 40 was less than 5%. Importantly, the final success for that number of GCs would require a ratio of the confidence θ random seeds to be successful (e.g. to have confidence $\theta = 0.6$ would required $6/10$ seeds to have distinct responses). Naturally we observed a strong relationship between 1. and 2. (Extended Data Fig. S17c), since the ability to successfully discriminate at the readout level depends on discrimination at the hidden GC level. In these experiments we selected $\theta = 0.9$.

Data and code availability

We used the PyTorch library for all neural network models. The code and respective simulated data used for our experiments is available at https://github.com/neuralml/ccLoops.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Supplementary Information

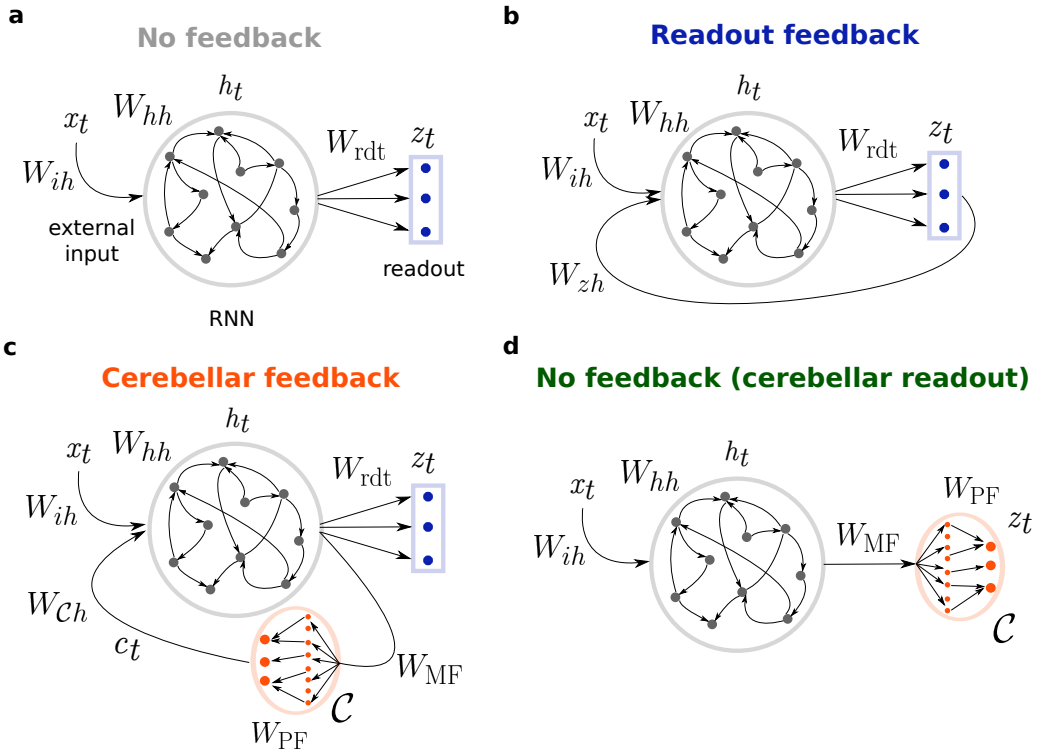

Figure S1. Different model architectures (extension of Fig. 1). **a**, *No feedback*; temporal input is fed to a cortical RNN (grey) and a linear readout layer (blue) produces the final model output. **b**, *Readout feedback*; now there is a feedback loop in which the RNN also receives readout predictions as extra input [12,14]. **c**, *Cerebellar feedback*; a copy of RNN activity is sent to a distinct but connected cerebellar network $\mathcal{C}$, which then returns its predictions back to the RNN as extra input. **d**, *No feedback with cerebellar readout*; like **c** a cerebellar network is attached to the RNN, but now it is used directly as the final readout and there is no "cortico-cerebellar loop". Model activity and weight vectors are represented with the same notation as Eqs. 1 and 2 (see also Table S1).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

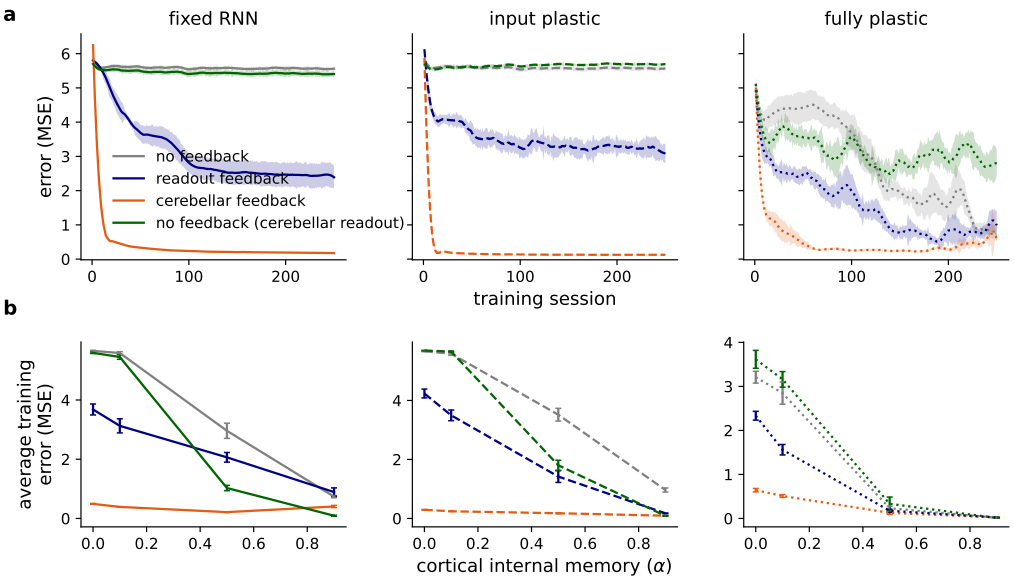

Figure S2. Model learning in the line drawing task. a, Training curves (cortical internal memory $\alpha = 0.1$) for the different models with fixed (left), input plastic (middle) and fully plastic (right) RNN. Green denotes the model where no feedback is applied to the RNN but the readout network (usually linear) now has the same architecture as the cerebellar network. **b**, Average error over training across different cortical internal memory α.

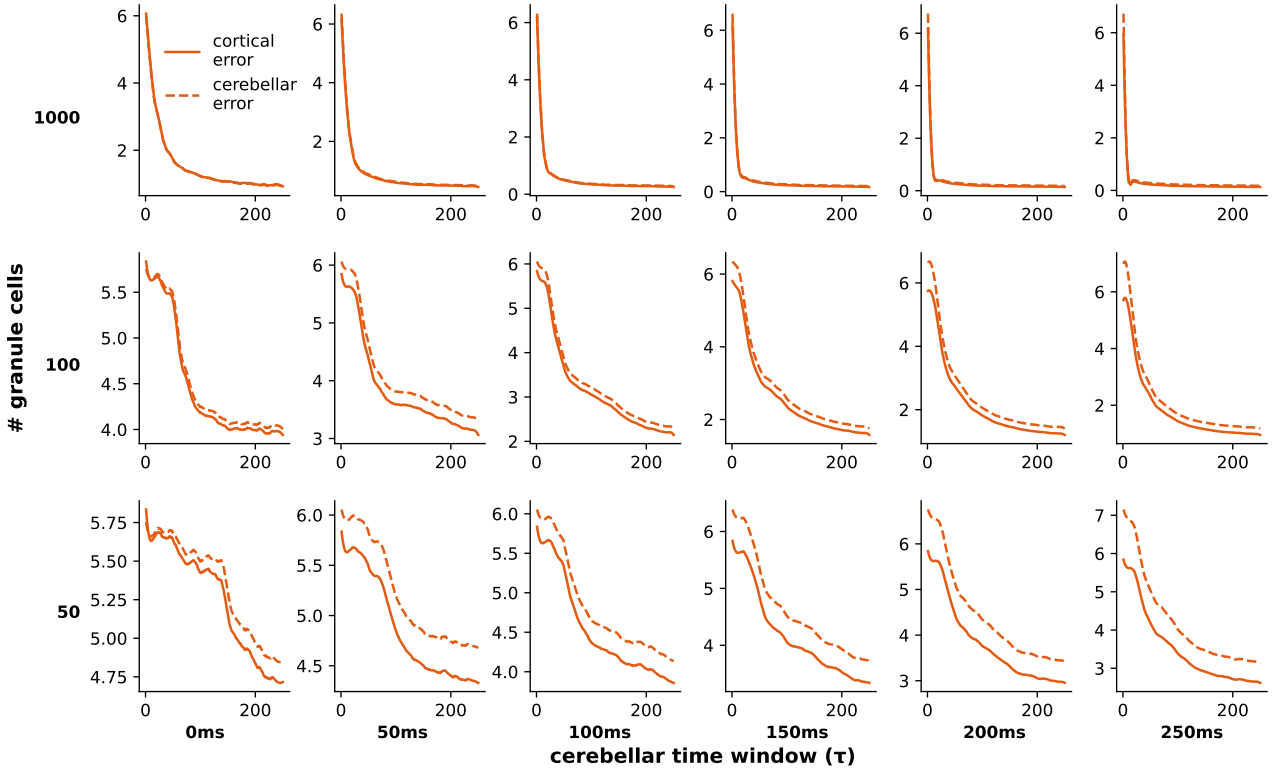

Figure S3. Training curves over different cerebellar parameters. We show the learning curves for the cortical network (solid line) and cerebellar network (dotted line) for the line drawing task. On each miniplot the x-axis represents the training session and y-axis the mean-squared error.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

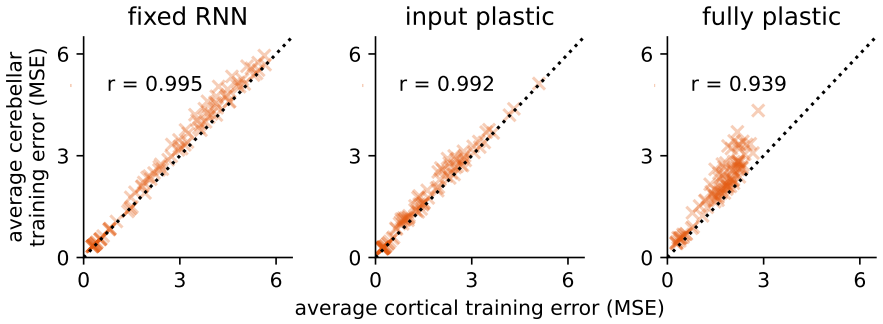

Figure S4. Integrated errors for the cortical network against the cerebellar network for the line drawing task. Each point denotes a specific cerebellar parameter configuration (1 of the 18 in Extended Data Fig. S3) and initialisation seed (1 of 5). *r* denotes the Pearson correlation.

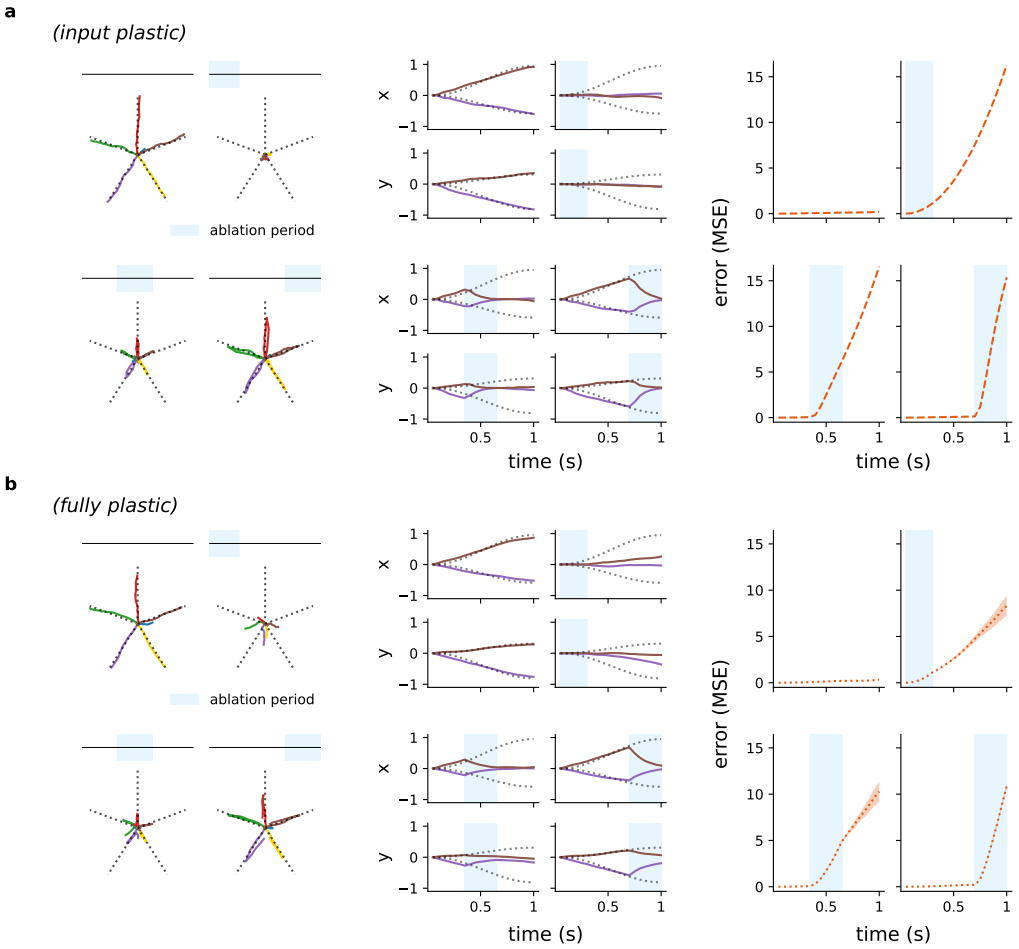

Figure S5. Ablation results for the input and fully plastic RNN. Model output (left, middle) and error (right) for example line drawing input under cerebellar ablation for an **a**, input plastic and **b**, fully plastic RNN. For the corresponding fixed RNN case see Fig. 2h-j.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

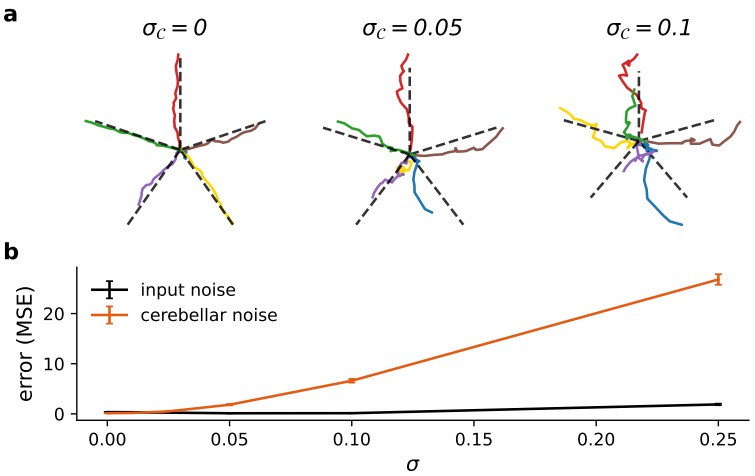

Figure S6. Cerebellar noise induces ataxic-like impairments. **a**, Output for trained models on the line drawing task under different levels of cerebellar noise $\mathbf{c}_t^{\text{noise}} = \mathbf{c}_t + \xi_t$ with $\xi_t \sim \mathcal{N}(\mathbf{0}; \sigma_c^2 \mathbf{I})$; blue output is for "no go" cue (where model is trained to remain at zero). **b**, Model error under various degrees of input and cerebellar noise.

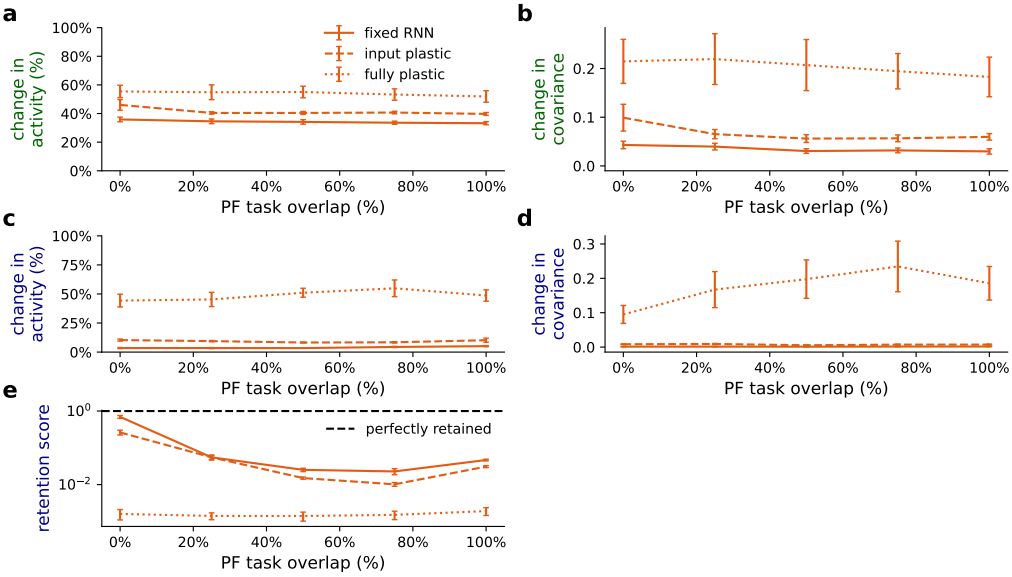

Figure S7. Multi-task learning and switching across different levels of parallel fibre (PF) overlap. **a**, **b**, Change in **a** activity and **b** covariance in RNN population between the line-drawing task 1 (baseline) and task 2 which is a curl-field variant. (**C**, **D**) Change in **c** activity and **d** covariance in RNN population between task 1 (baseline) and task 1 (post re-learning) after switching back from task 2. **e**, Task 1 retention score, which is computed as the error of task 1 during baseline over the error at the first trial after switching back to task 1.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

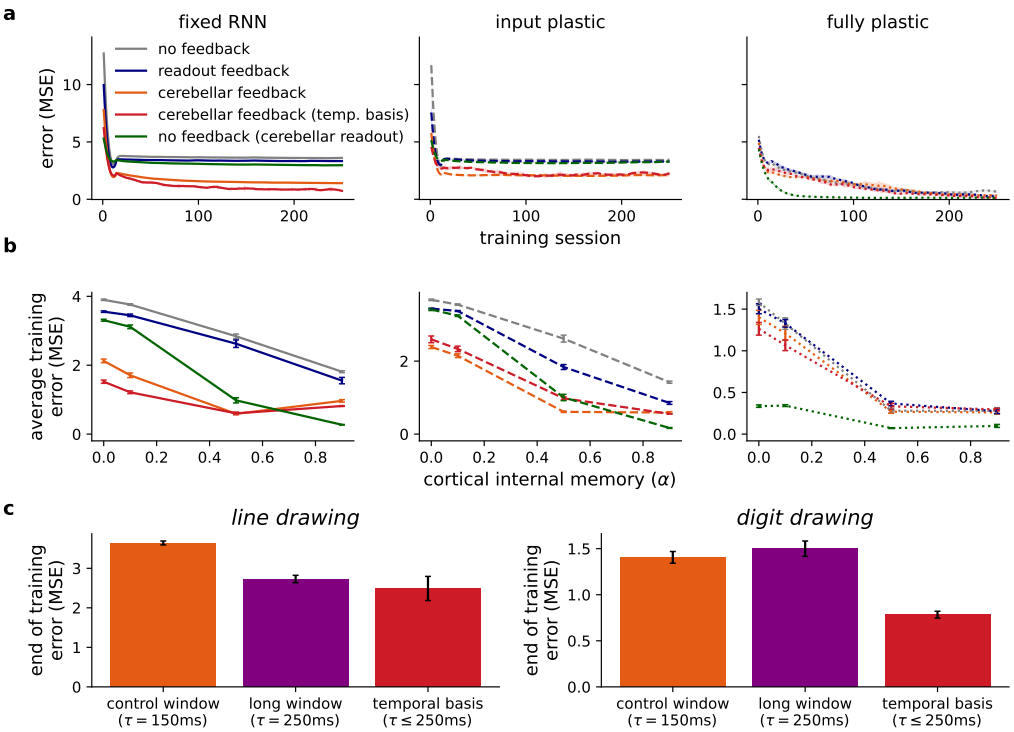

Figure S8. Model learning in the digit drawing task. a, Training curves (cortical internal memory $\alpha = 0.1$) for the different models with a fixed (left), input plastic (middle) and fully plastic (right) RNN plasticity assumptions. Green denotes the model where no feedback is applied to the RNN but the readout network (usually linear) now has the same architecture as the cerebellar network. **b**, Average error over training across different cortical internal memory α. **c**, Model error (fixed RNN; $\alpha = 0.1$) at the end of training (averaged over last 10 training sessions) for different cerebellar time windows for (left) line drawing task (cf. Fig. 2) and (right) digit drawing task.

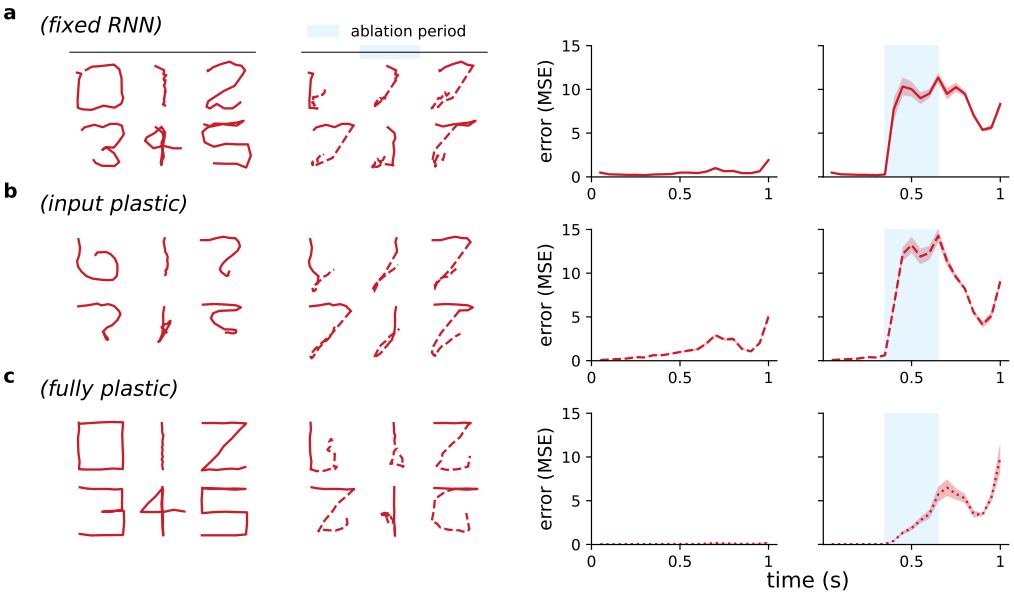

Figure S9. Ablation results for the digit drawing task. Model output (left) and error (right) for digit drawing input under cerebellar ablation for a **a**, fixed, **b**, input plastic, and **c**, fully plastic RNN. Model output shown after cerebellar ablation with dotted lines.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

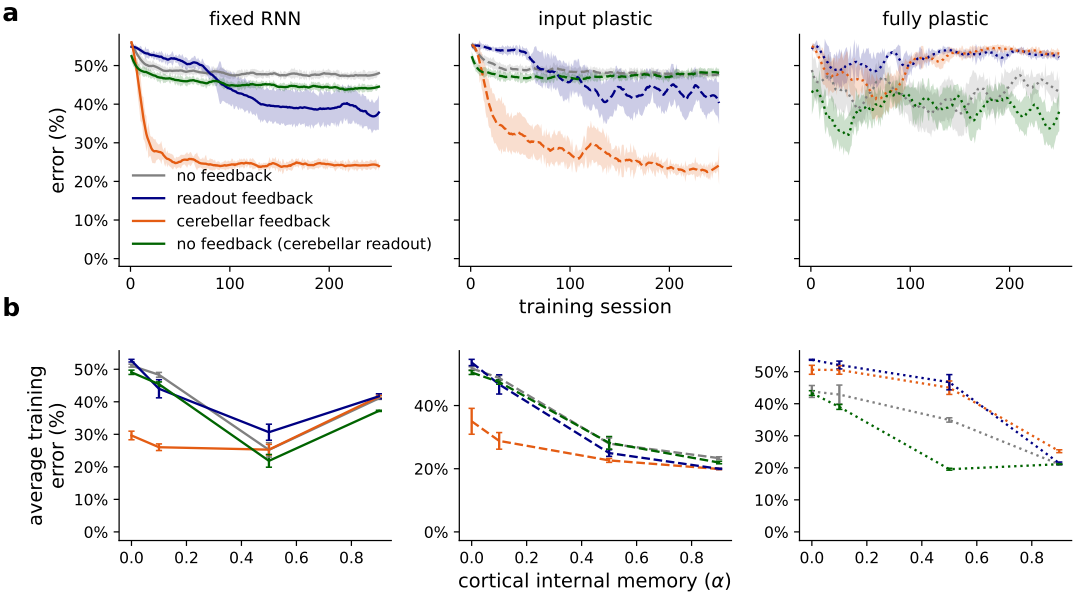

Figure S10. Model learning in the evidence accumulation task. **a**, Training curves (cortical internal memory $\alpha = 0.1$) for the different models with a fixed RNN (left), input plastic (middle) and fully plastic (right) RNN. Green denotes the model where no feedback is applied to the RNN but the readout network (usually linear) now has the same architecture as the cerebellar network. **b**, Average error over training across different levels of cortical internal memory α.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

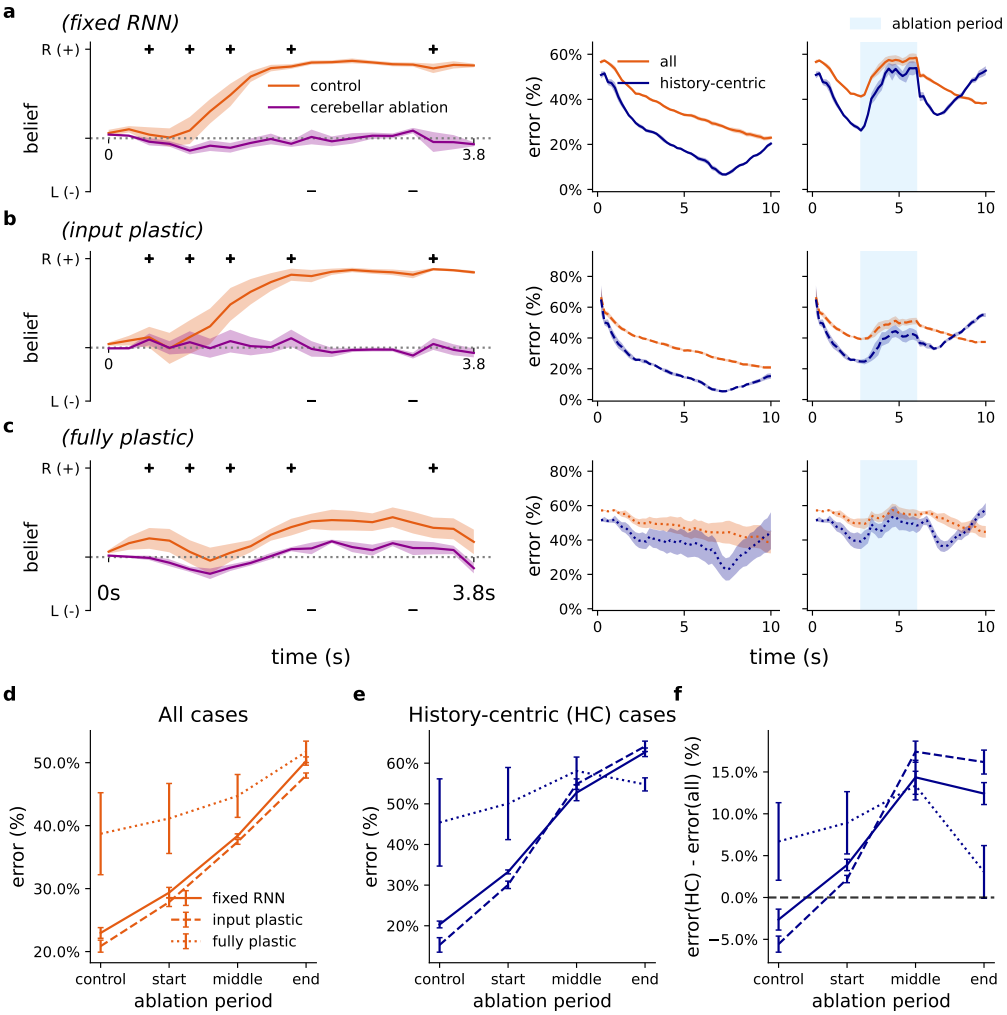

Figure S11. Additional cerebellar ablation results for evidence accumulation task. **a**-**c**, Model output (left) and error (right) with and without cerebellar ablation (model output shows full cerebellar ablation case) for **a** fixed, **b** input plastic, and **c** fully plastic RNN. **d**, The error for different ablation periods across these RNN plasticity conditions over all test examples. **e**, The error for different ablation periods across different RNN plasticity conditions, but only over "history-centric" inputs. **f**, The difference between **d** and **e**.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

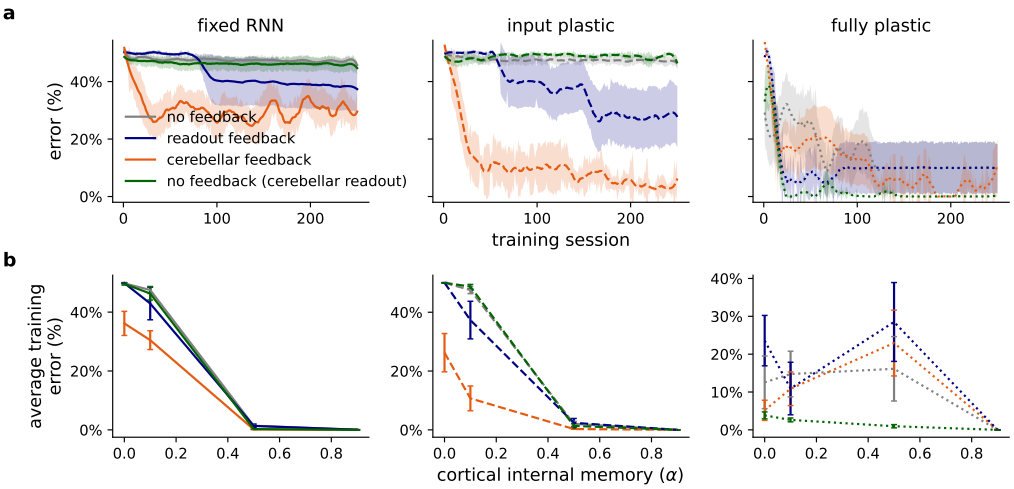

Figure S12. Learning curves in the delayed association task. **a**, Training curves (cortical internal memory $\alpha = 0.1$) for the different models with a fixed (left), input plastic (middle) and fully plastic (right) RNN. Green denotes the model where no feedback is applied to the RNN but the readout network (usually linear) now has the same architecture as the cerebellar network. **b**, Average error over training across different cortical internal memory α.

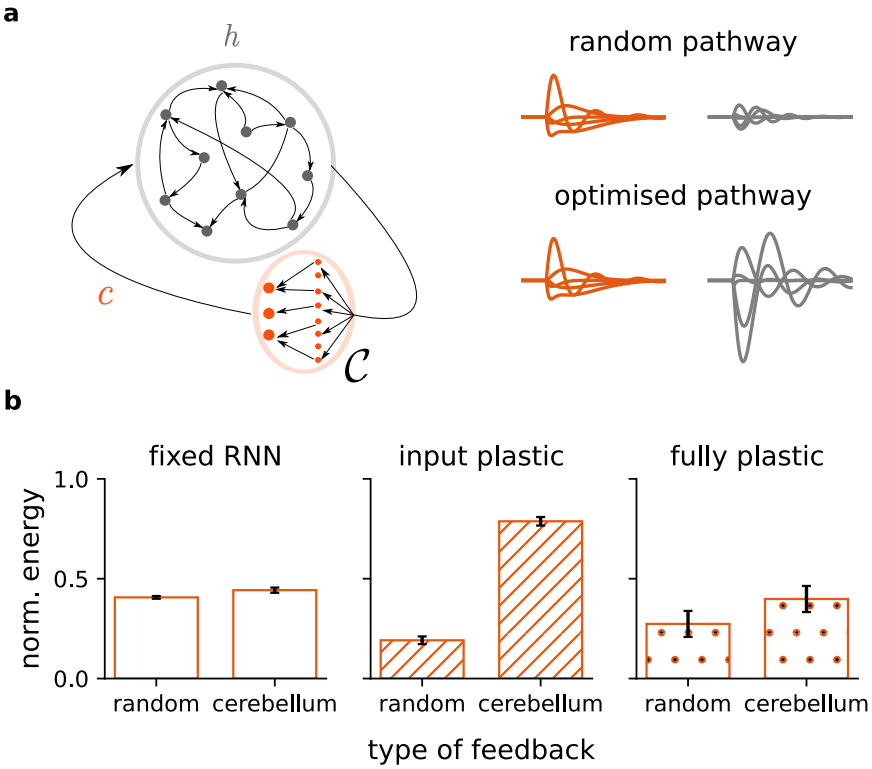

Figure S13. A control-theoretic perspective of the cortico-cerebellar loop. **a**, Illustrative schematic of cerebellar (orange) and cortical (grey) activities. Depending on the cerebello-cortical connectivity $W_{\mathcal{C}h}$, the same cerebellar output **c** might suppress (top right) or amplify (bottom right) RNN trajectories. **b**, The energy (see Methods) generated by random and cerebellar feedback for models trained with varying degrees of plasticity in the delayed association task (Fig. 6). The energy is normalised by the maximum possible energy generated by inputs that achieve the greatest cortical response (see Methods).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

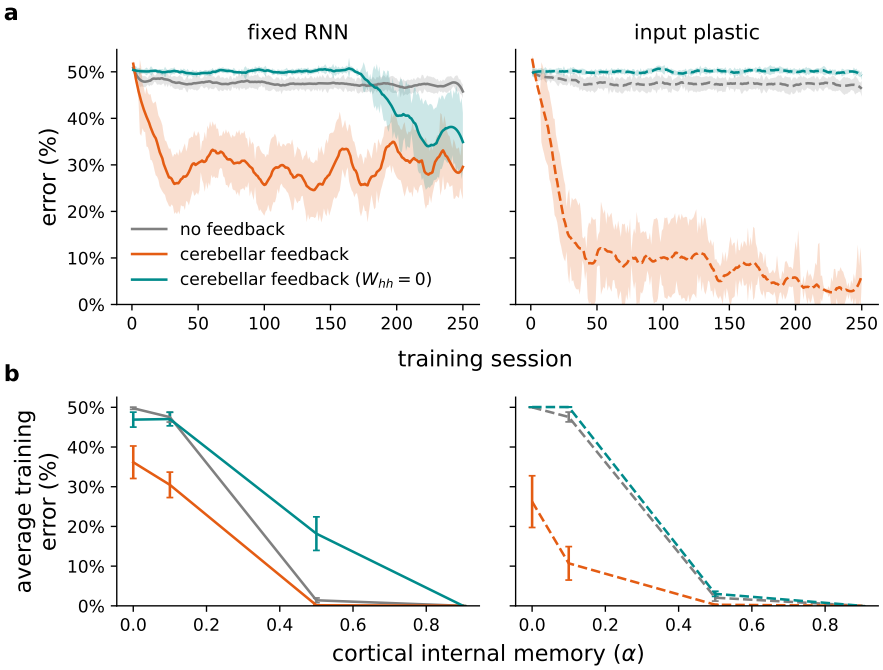

Figure S14. Cerebellar feedback with cortical recurrent connectivity is necessary to learn long-range temporal associations. **a**, Training curves (cortical internal memory $\alpha = 0.1$) with cerebellar feedback but zero recurrent weights ($W_{hh} = 0$) with a fixed (left) and input plastic (right) RNN **b**, Average error over training across different cortical internal memory α.

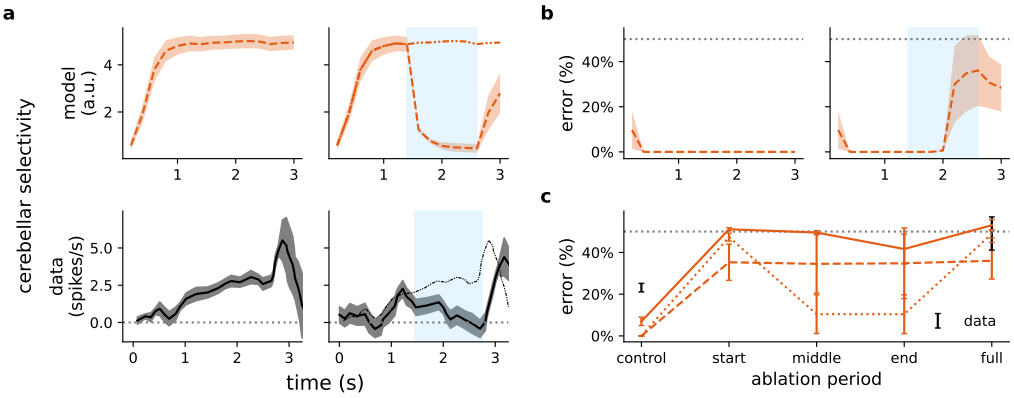

Figure S15. Cortical ablation in the delayed association task. **a**, Cue selectivity in the cerebellar network during the delay period without (left) and with cortical ablation (right; blue area denotes period of ablation; 75% of the cortical RNN neurons are silenced) conditions for example input in model (upper panels) and experimental data (lower panels) reproduced from Gao et al. [23]. **b**, (Cortical) model error during delay period with (left) and without (right) cortical ablation. **c**, Average error from cortical ablation at different periods during the task delay period and different degrees of plasticity. Experimental data shown in black.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

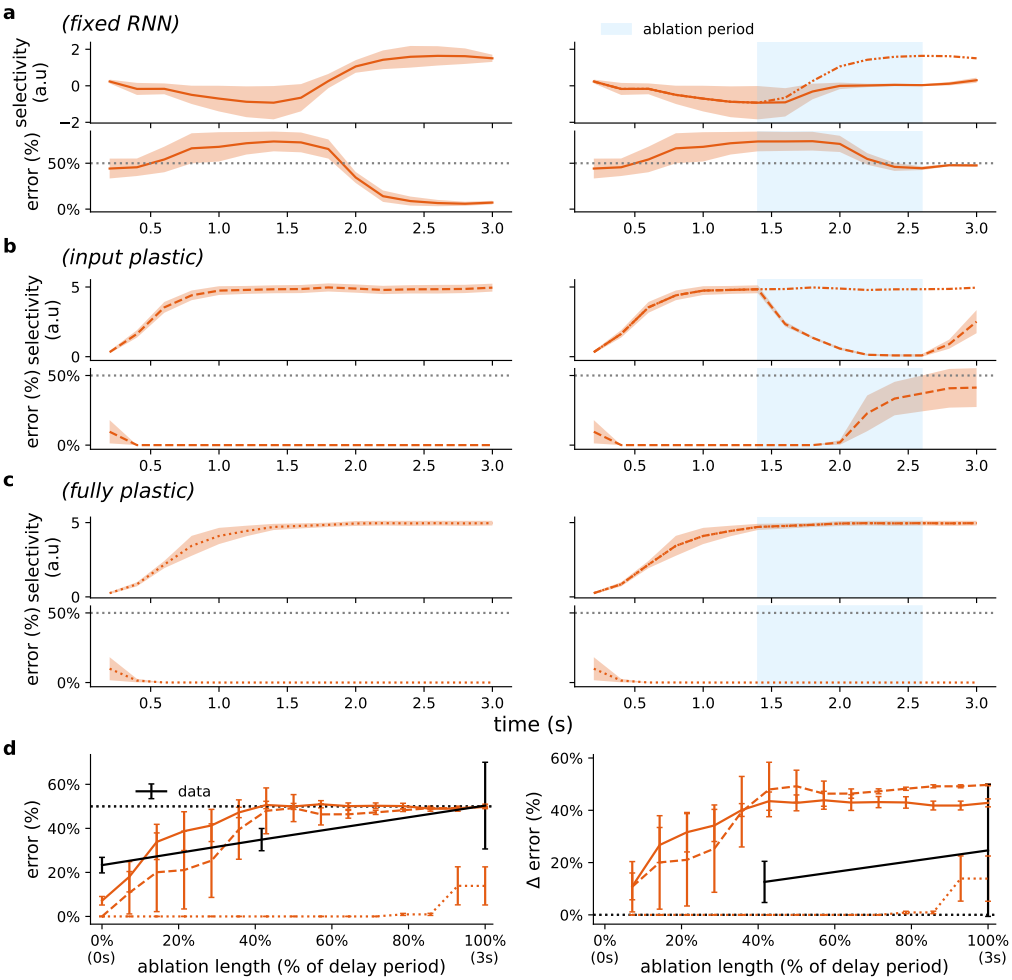

Figure S16. Additional cerebellar ablation results for the delayed association task. **a**-**c**, Model output (top) and error (bottom) for the delayed association task without (left) and with (right) cerebellar ablation with a **a** fixed, **b** input plastic, and **c** fully plastic RNN. Thin line after ablation shows control model. **d**, Model error as a function of ablation length (centred around the middle of the delay period). Experimental data reproduced from Gao et al. [23]. Dotted black line denotes chance.

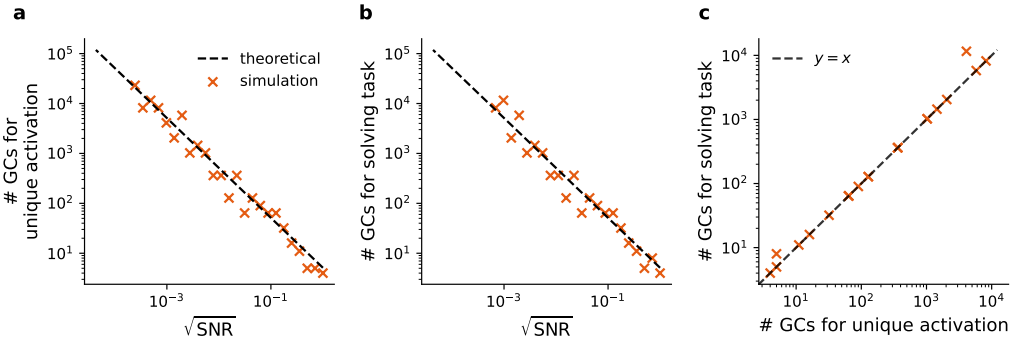

Figure S17. Analytical and numerical relationship between input signal-to-noise ratio (SNR) and number of granule cells (GCs) for **a**, distinct population encoding (Fig. 7c) and **b**, successful readout learning; $C/\sqrt{\text{SNR}}$ denotes the theoretical prediction (see early section on the cerebellum decoding fading memory). **c**, Tight coupling between **a** and **b**.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516257; this version posted November 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

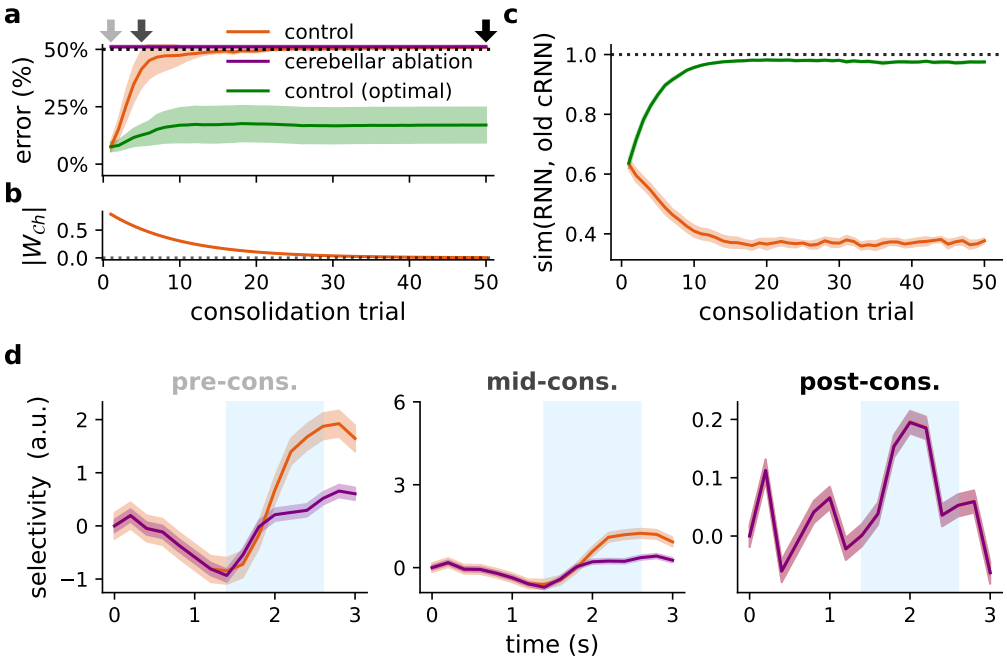

Figure S18. Cerebellar-to-cortical consolidation of the delayed association task with fixed RNN models. **a**, Accuracy of control and cerebellar ablation conditions (dotted line denotes chance) and the corresponding **b**, strength of the cerebellar-cortical pathway (W_{Ch}) over consolidation. Green denotes control condition with theoretically optimal learning rule. **c**, Cosine similarity between cortico-cortical input and total cortical input (i.e. cerebellar-cortical and cortico-cortical inputs) pre-consolidation. Similarity of the consolidation model is shown in orange and the optimal consolidation model in green. **d**, Model selectivity for example (external) input in control and cerebellar ablation conditions at different stages of the consolidation process; colour coded by arrow times in **a**.