

---

# Minimizing Control for Credit Assignment with Strong Feedback

---

Alexander Meulemans<sup>\*1</sup> Matilde Tristany Farinha<sup>\*1</sup> Maria Cervera<sup>\*1</sup>  
João Sacramento<sup>1</sup> Benjamin F. Grewe<sup>1</sup>

## Abstract

The success of deep learning ignited interest in whether the brain learns hierarchical representations using gradient-based learning. However, current biologically plausible methods for gradient-based credit assignment in deep neural networks need infinitesimally small feedback signals, which is problematic in biologically realistic noisy environments and at odds with experimental evidence in neuroscience showing that top-down feedback can significantly influence neural activity. Building upon deep feedback control (DFC), a recently proposed credit assignment method, we combine *strong* feedback influences on neural activity with gradient-based learning and show that this naturally leads to a novel view on neural network optimization. Instead of gradually changing the network weights towards configurations with low output loss, weight updates gradually minimize the amount of feedback required from a controller that drives the network to the supervised output label. Moreover, we show that the use of strong feedback in DFC allows learning forward and feedback connections simultaneously, using learning rules fully local in space and time. We complement our theoretical results with experiments on standard computer-vision benchmarks, showing competitive performance to backpropagation as well as robustness to noise. Overall, our work presents a fundamentally novel view of learning as control minimization, while sidestepping biologically unrealistic assumptions.

## 1. Introduction

The error backpropagation (BP) method (Rumelhart et al., 1986; Werbos, 1982) has emerged as the method of choice

---

<sup>\*</sup>Equal contribution <sup>1</sup>Institute of Neuroinformatics, University of Zürich and ETH Zürich, Switzerland. Correspondence to: Alexander Meulemans <ameulema@ethz.ch>.

for training deep artificial neural networks, due to its efficient computation of gradients, which is crucial for learning in high-dimensional parameter spaces (LeCun et al., 2015). Although biological networks in the neocortex form similar hierarchies (Yamins & DiCarlo, 2016), it is not clear yet if they follow similar principles to implement credit assignment (CA), i.e., determining how to change synaptic connectivity strengths to get closer to some desired output. Several key properties of the BP method are highly incompatible with cortical networks (Crick, 1989; Lillicrap et al., 2020). For example, synaptic plasticity in biological neurons is local in space and time and tightly coupled to neural activity. By contrast, artificial neural networks trained with BP process activity and weight update signals separately and in distinct phases while requiring exactly the same synaptic weights in the forward and feedback pathways.

To couple synaptic plasticity directly to neural activity, recent work designed cortical network models that propagate CA signals by leveraging the network dynamics (Scellier & Bengio, 2017; Whittington & Bogacz, 2017; Sacramento et al., 2018; Meulemans et al., 2021) or using multiplexed neural signals (Payeur et al., 2021). However, these methods require tightly coordinated plasticity mechanisms, possibly through the use of distinct training phases, and it remains unclear whether cortical networks can exhibit these levels of coordination. Moreover, to approximate gradient-based optimization these methods consider the *weak-feedback* limit where, to avoid interference, feedback cannot significantly alter the feedforward processing of the network. This weak-feedback assumption is problematic in biologically realistic noisy environments, and is in stark contrast with biological observations showing that top-down feedback can significantly affect forward processing (Jordan & Keller, 2020; Keller et al., 2020).

To date, several studies used strong feedback control signals in combination with error-based learning rules to drive the network activity to track a reference output (Gilra & Gerstner, 2017; Alemi et al., 2018; Denève et al., 2017; Bourdoukan & Deneve, 2015). Although these models use strong feedback influences on neural activity, they have only been successfully applied to train single-layer recurrent neural networks with fixed feedback weights and it remains unclear whether they can be extended for training

deep neural networks. Furthermore, even though these methods are capable of reducing the output loss during training, it remains an open question whether they extract gradient information to enable efficient learning. Hence, gradient-based CA theories for training deep neural networks with strong feedback influences and without the need for tightly coordinated plasticity mechanisms are still lacking.

Building upon Deep Feedback Control (DFC) (Meulemans et al., 2021), we introduce Strong-DFC, a gradient-based approach to CA in deep neural networks that makes use of strong feedback influences on neural activity. Like DFC, we use a feedback controller to drive the network to a desired output target and update the network weights with a learning rule fully local in space and time which leverages the dynamic change in neural activity. Unlike DFC, we allow for strong feedback influences by taking the supervised label as the desired output target instead of weakly nudging the current network output towards the supervised label. This new setting results in significant feedback influence on the neural activities, as the initial network output is far from the supervised label at the beginning of training.

Interestingly, because the output targets remain invariant throughout training, Strong-DFC naturally leads to a fundamentally different perspective on neural network optimization. As the controller drives the network to the supervised output labels at each point in training, the goal of optimizing the network parameters shifts from minimizing the supervised loss towards minimizing the amount of feedback control needed to reach the supervised output labels. To formalize this *minimizing control* perspective we introduce a surrogate loss that quantifies the amount of control feedback, and show that, under flexible conditions on the feedback connectivity, the Strong-DFC updates follow its negative gradient. Intriguingly, Strong-DFC enables us to learn the forward and feedback weights simultaneously, overcoming the two-phase requirement of DFC, while inheriting the close connection to recently proposed dendritic compartment models of cortical pyramidal neurons equipped with voltage-dependent plasticity rules (Sacramento et al., 2018). Finally, we complement our theoretical findings with experimental results on standard computer vision benchmarks, showing that Strong-DFC provides principled CA in deep neural networks using strong feedback influences without the need for tightly coordinated plasticity mechanisms.

## 2. Background

We revisit DFC (Meulemans et al., 2021) to investigate CA with strong feedback influence on neural activities. The control framework used by DFC was designed for feedforward networks with continuous dynamics. More specifically, DFC considers a multilayer network with the following dynamics (for  $1 \leq i \leq L$ ):

$$\tau_v \dot{\mathbf{v}}_i(t) = -\mathbf{v}_i(t) + W_i \phi(\mathbf{v}_{i-1}(t)) + Q_i \mathbf{u}(t), \quad (1)$$

where  $\mathbf{v}_i$  is a vector of pre-nonlinearity activations in layer  $i$ ,  $\dot{\mathbf{v}}_i$  its time derivative,  $W_i$  the forward weights,  $\phi$  a smooth nonlinear activation function,  $Q_i$  the feedback weights, and  $\tau_v$  the neural time constant. We define  $\mathbf{r}_i = \phi(\mathbf{v}_i)$  as the post-nonlinearity activity of layer  $i$ . Note that without feedback influences ( $\mathbf{u} = 0$ ), the steady state of the above dynamics corresponds to a feedforward neural network, as the input  $\mathbf{r}_0$  remains constant over time.

The feedback signal  $\mathbf{u}$  is computed by a proportional integral feedback controller, using the output error  $\mathbf{e}(t) = \mathbf{r}_L - \mathbf{r}_L(t)$ , and serves to push the network to a desired output target  $\mathbf{r}_L$ :

$$\mathbf{u}(t) = \mathbf{u}^{\text{int}}(t) + k\mathbf{e}(t), \quad \tau_u \dot{\mathbf{u}}^{\text{int}}(t) = \mathbf{e}(t) - \alpha \mathbf{u}^{\text{int}}(t), \quad (2)$$

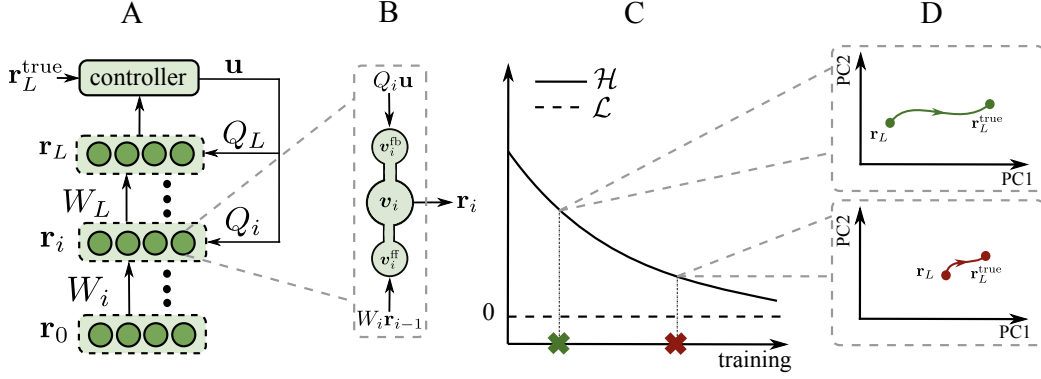
where  $k$  is the proportional control constant and  $\alpha$  the leakage constant limiting the magnitude of  $\mathbf{u}^{\text{int}}$ . DFC weakly nudges the network output in the direction of lower loss, since the output target  $\mathbf{r}_L$  is defined as:

$$\mathbf{r}_L = \mathbf{r}_L - \lambda \left. \frac{\partial \mathcal{L}}{\partial \mathbf{r}_L} \right|_{\mathbf{r}_L = \mathbf{r}_L^-}^T, \quad (3)$$

with  $\mathbf{r}_L$  the steady state of the output without the presence of a controller,  $\mathcal{L}$  the loss function, and  $\lambda$  the nudging strength. The theoretical results linking DFC to a variant of Gauss-Newton optimization consider the *weak-feedback* regime of  $\lambda \rightarrow 0$  (Meulemans et al., 2021). Note that we cannot enforce *strong feedback* in this theoretical framework by keeping the small nudging factor  $\lambda \rightarrow 0$  and making  $Q_i$  arbitrarily large, as the limit considered by Meulemans et al. (2021) relies on vanishing feedback influence on neural activity, similar to other seminal work in dynamic CA methods that approximate gradient-based optimization (Sacramento et al., 2018; Scellier & Bengio, 2017; Whittington & Bogacz, 2017). DFC then updates the feedforward weights with the following plasticity rule that is fully local in space and time:

$$\tau_W \dot{W}_i(t) = (\phi(\mathbf{v}_i(t)) - \phi(W_i \mathbf{r}_{i-1}(t))) \mathbf{r}_{i-1}(t)^T. \quad (4)$$

Using a multi-compartment model of a neuron (Fig. 1B), one can interpret this learning rule as a delta rule that uses the difference between the outgoing firing rate of the neuron  $\phi(\mathbf{v}_i)$  and the estimated firing rate based only on the feedforward input of the neuron  $\phi(\mathbf{v}_i^{\text{ff}}) = \phi(W_i \mathbf{v}_{i-1})$  (Urbanczik & Senn, 2014). Standard DFC learns its feedback weights  $Q_i$  in a separate ‘sleep’ phase with a simple anti-Hebbian learning rule, making use of the feedback compartment of the multi-compartment neuron,  $\mathbf{v}_i^{\text{fb}} = Q_i \mathbf{u}$ . Note that DFC assumes a separation of timescales between the plasticity



**Figure 1. Strong-DFC trains deep neural networks by minimizing control.** (A) Schematic of a feedforward network trained with Strong-DFC, where feedback weights  $Q$  are uncoupled from forward weights  $W$ , and used to carry a control signal  $\mathbf{u}$  from the controller to all hidden neurons. (B) Neurons are multi-compartmental and allow integrating separate forward drive  $v_i^{\text{ff}}$  and feedback control  $v_i^{\text{fb}}$  signals. (C) Since the supervised target  $\mathbf{r}_L^{\text{true}}$  is always reached whenever the controller is on, the training loss  $\mathcal{L}$  is equal to zero throughout, and training can be better understood as the minimization of a surrogate loss  $\mathcal{H}$  describing the amount of control required. (D) Early in training (top) the controller feedback strongly influences the neural activity to drive the network output to the desired target, and training minimizes the amount of control required (bottom). Figure A-B adapted from Meulemans et al. (2021).

dynamics (4) and the network dynamics (1-2), such that the weights  $W_i$  and  $Q_i$  can be considered constant on the timescale of the network-controller interactions.

### 3. Training neural networks with strong feedback by minimizing control

Here, we introduce our *minimizing control* framework that allows for gradient-based CA with strong feedback influences in deep neural networks. We use the feedback controller (2) to strongly drive the network towards the supervised output label at each point during training. Consequently, the focus of optimizing the network parameters shifts from minimizing the training loss  $\mathcal{L}$  towards minimizing the amount of feedback control needed to reach the supervised output label. We quantify the magnitude of the control input with a surrogate loss  $\mathcal{H}$ , and show that under flexible constraints on the feedback connectivity, the Strong-DFC updates perform gradient descent on this new loss function. Finally, we show that by minimizing  $\mathcal{H}$ , we implicitly minimize  $\mathcal{L}$  as well. Throughout this section, we assume the network is stable, on which we elaborate at the end of the section. Although the theoretical results shown here are tailored towards the proportional integral controller (2), these can easily be extended to other controller types such as a pure proportional controller (App. A.5).

**Strong-DFC.** We adapt DFC to allow for strong feedback influence on neural activity by setting the output target  $\mathbf{r}_L$  to the supervised output label  $\mathbf{r}_L^{\text{true}}$ , instead of weakly nudging it towards values of lower loss (3) (see Fig. 1). This new setting results in strong feedback influence, as the network output without feedback is far from the supervised labels at the beginning of training. Furthermore, as the existing

theory on DFC relies on the assumption that the feedback only infinitesimally changes the network activity, it is no longer valid in this new setting of DFC with strong feedback (Strong-DFC).

**Learning as control minimization.** Here, we propose a novel view on neural network optimization that justifies strong feedback influences on neural activity. In Strong-DFC, at every stage of training the feedback controller drives the network to the supervised label  $\mathbf{r}_L^{\text{true}}$ , resulting in zero training loss at steady state. Hence, instead of gradually changing the network weights towards configurations with low output loss, the goal of optimizing the network parameters is now to gradually reduce the amount of feedback required from the controller to reach the supervised output label (see Fig. 1C-D). At the individual neuron level, one can interpret this new perspective on optimization as each neuron minimizing its feedback input by altering its feedforward weights.

We formalize this notion of *minimizing control* by introducing a surrogate loss function representing the magnitude of the feedback control input:

$$\mathcal{H} = \frac{1}{2} \sum_b \|Q \mathbf{u}_{\text{ss}}^{(b)}\|_2^2, \quad (5)$$

with  $Q = [Q_1^T, \dots, Q_L^T]^T$  the concatenated feedback weights and  $\mathbf{u}_{\text{ss}}^{(b)}$  the steady-state control feedback of a data point  $b$ . The control feedback  $\mathbf{u}_{\text{ss}}$  depends on the feedforward weights  $W_i$  through the network and controller dynamics (Eq. 1-2). Using the implicit function theorem, in Theorem 1 we compute the total derivative of  $\mathcal{H}$  w.r.t.  $W_i$  and show that, under flexible constraints on the feedback weights, it results in a local update in space and time that aligns with a

simplified version of the Strong-DFC weight updates (full proof in App. A.1).

**Theorem 1.** *Assuming  $J_{ss}Q + \tilde{\alpha}I$  is invertible, the total derivative of  $\mathcal{H}$  w.r.t.  $W_i$  is given by:*

$$\frac{d\mathcal{H}}{dW_i^T} = - \sum_b J_{i,ss}^{(b)T} \left( J_{ss}^{(b)}Q + \tilde{\alpha}I \right)^T Q^T Q \mathbf{u}_{ss}^{(b)} \mathbf{r}_{i,1,ss}^{(b)T}, \quad (6)$$

with  $J_{i,ss}$ ,  $\frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_i}$  and  $J_{ss}$ ,  $\left[ \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1}, \dots, \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L} \right]$ , both evaluated at steady state, and  $\tilde{\alpha} = \frac{\alpha}{1+k\alpha}$ .

Furthermore, assuming Condition 1 holds, the network is stable,  $\alpha \rightarrow 0$ , and  $J_{ss}Q + \tilde{\alpha}I$  is invertible, then the following steady-state (ss) updates for the forward weights, averaged over the batch samples,

$$\Delta W_{i,ss} = \eta (\mathbf{v}_{i,ss} - \mathbf{v}_{i,ss}^{\text{ff}}) \mathbf{r}_{i,1,ss}^T, \quad (7)$$

with  $\eta$  the stepsize, align with the negative gradient  $-\frac{d\mathcal{H}}{dW_i}$ .

**Condition 1.** The column space of the concatenated feedback weights  $Q$  is equal to the row space of  $J_{ss}$ ,  $\left[ \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1}, \dots, \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L} \right]_{\mathbf{v}_i = \mathbf{v}_{i,ss}}$ , the network Jacobian at steady state.

Theorem 1 links Strong-DFC to gradient descent on a surrogate loss function  $\mathcal{H}$ . Although the steady-state updates (7) of Theorem 1 are not exactly equal to the DFC updates (4), previous work showed that it is a good approximation (see App. A.3-A.4 and Meulemans et al., 2021).

**Flexible feedback connectivity constraint.** Condition 1 is a flexible connectivity constraint on the feedback weights that represents the requirement that the feedback controller influences the network activity in the most efficient manner (see App. A.2). More specifically, if Condition 1 is not satisfied, the feedback input  $Q\mathbf{u}_{ss}$  is partially lost in the nullspace of the network and hence the same influence on the network output could be reached with a smaller feedback input. Interestingly, Condition 1 generalizes the feedback requirements of commonly used learning rules for approximating gradient-based optimization methods. For example, backpropagation and its approximations require  $Q = J^T$  (Rumelhart et al., 1986; Lansdell et al., 2020; Akrouf et al., 2019) and common variants of target propagation require the pseudoinverse  $Q = J^\dagger$  (Meulemans et al., 2020), both satisfying Condition 1. More generally, Condition 1 implies that the feedback weights of Strong-DFC do not necessarily need to align with the forward pathway or its inverse to provide gradient-based CA by following the negative gradient of  $\mathcal{H}$ . Remarkably, the fundamentally distinct theoretical framework of standard DFC gives rise to a similar condition on  $Q$ , however with  $J$  evaluated at the feedforward activities without feedback present instead of the steady-state activities (Meulemans et al., 2021).

**Minimizing  $\mathcal{H}$  leads to minimizing  $\mathcal{L}$ .** Theorem 1 links Strong-DFC to gradient descent on the surrogate loss function  $\mathcal{H}$ . Hence, to show that Strong-DFC minimizes the original loss function  $\mathcal{L}$ , we need to connect  $\mathcal{H}$  to  $\mathcal{L}$ . Proposition 2 shows that if  $\mathcal{H}$  is minimized to zero,  $\mathcal{L}$  is minimized to zero as well, which is trivial to prove as  $\mathcal{H} = 0$  translates to all training samples reaching exactly their supervised output label without help from the controller.

**Proposition 2.** *For all  $\mathcal{L}$  that obey  $\mathcal{L}(\mathbf{x}, \mathbf{x}) = 0$ , we have that*

$$\mathcal{H} = \sum_b \|Q\mathbf{u}_{ss}^{(b)}\|_2^2 = 0 \iff \sum_b \mathcal{L}(\mathbf{r}_L^{(b)}, \mathbf{r}_L^{\text{true},(b)}) = 0,$$

with  $\mathbf{r}_L$  the steady state of the output without the presence of a controller.

Recent work in the field of deep learning has shown that, for overparameterized networks, it is possible to minimize the training loss  $\mathcal{L}$  to zero while reaching good generalization (Belkin et al., 2019). We show empirically in Section 5 that Strong-DFC can minimize  $\mathcal{H}$ , and hence  $\mathcal{L}$  to zero, thereby validating the minimizing control framework for optimizing neural networks.

**Stability of Strong-DFC.** So far, we assumed that the network and controller dynamics reach their steady state. However, this is not a given, as the feedback interaction could lead to unstable loops causing the network to diverge. Adapting the results of Meulemans et al. (2021), we investigate the local stability of a simplified version of Strong-DFC that uses integral control and assumes a separation of timescales  $\tau_v \ll \tau_u$ . At steady state, the Jacobian matrix of this dynamical system is equal to  $-(J_{ss}Q + \alpha I)$ , resulting in the following stability condition on the feedback weights (see App. B).

**Condition 2.** Given the network Jacobian evaluated at steady state  $J_{ss}$ , the real parts of the eigenvalues of  $-J_{ss}Q$  are all below  $\alpha$ .

In App. B, we compute the Jacobian matrix of the full dynamical system representing Strong-DFC with proportional integral control and with no assumptions on a separation of timescales. However, a local stability analysis on this system does not reveal clearly interpretable stability conditions for the feedback weights. Therefore, we monitor instead the stability of the full system empirically.

## 4. A single-phase learning scheme for forward and feedback weights

Conditions 1 and 2 highlight the importance of the feedback weights for learning and stability, respectively. As training causes the forward weights and thus the network Jacobian



$J_{ss}$  to change, the set of feedback connectivity patterns satisfying Conditions 1 and 2 shift during learning as well. To keep the feedback weights compatible with the changing network, we learn the feedback weights together with the forward weights.

Here, we introduce plasticity dynamics for the feedback weights that make it possible to update both forward and feedback weights simultaneously in a single phase, in contrast to standard DFC and other methods that need two separate phases. Our new feedback plasticity dynamics leverage two essential properties of Strong-DFC. (i) The strong feedback influences make it possible to have noise in the dynamics while learning the forward weights, while keeping a good signal-to-noise ratio of the learning signals. (ii) The conditions on the feedback weights for learning and stability required for our minimizing control framework consider the network activity at steady state. Hence, the feedback weights should be learned at the same steady state used for training the forward weights. Leveraging the noise intrinsically available to neurons, we use a simple anti-Hebbian learning rule for the feedback weights to drive them towards satisfying Conditions 1 and 2.

**Noisy neural dynamics.** Our feedback weight plasticity dynamics make use of noise injected in the neural dynamics, which we model as follows:

$$\tau_v \dot{\mathbf{v}}_i(t) = -\mathbf{v}_i(t) + W_i \phi(\mathbf{v}_{i-1}(t)) + Q_i \mathbf{u}(t) + \sigma \epsilon_i(t), \quad (8)$$

where the noise  $\epsilon$  represents exponentially filtered white noise with time constant  $\tau_\epsilon$ , i.e., an Ornstein-Uhlenbeck process, and  $\sigma$  the standard deviation of the noise. Note that such Ornstein-Uhlenbeck processes can, for example, be obtained in the limit of many independent Poisson inputs with infinitesimal synaptic weights (Gerstner et al., 2014).

**An anti-Hebbian learning rule for  $Q$ .** The noise fluctuations that propagate through the network carry information about the network Jacobian  $J$  and we can extract this information by correlating the noise fluctuations at the output with the noise fluctuations at each layer. Inspired by the feedback weight learning of standard DFC (Meulemans et al., 2021), we instantiate this correlation technique with the following simple anti-Hebbian plasticity rule:

$$\tau_Q \dot{Q}_i(t) = - \left(1 + \frac{\tau_v}{\tau_\epsilon}\right)^L \mathbf{v}_i^{\text{fb}}(t) \mathbf{u}_{\text{hp}}(t)^T - \beta Q_i. \quad (9)$$

The feedback compartment  $\mathbf{v}_i^{\text{fb}}$  contains a part of the injected noise,  $\mathbf{v}_i^{\text{fb}} = Q_i \mathbf{u} + \sigma \epsilon_i^{\text{fb}}$ , and the pre-synaptic plasticity signal  $\mathbf{u}_{\text{hp}}$  is a high-pass filtered version of the incoming control signal  $\mathbf{u}$ , which extracts the output noise

fluctuations.<sup>1</sup> Our feedback weight plasticity dynamics (9) introduce two important novelties with respect to DFC. (i) The high-pass filtering of  $\mathbf{u}$  makes it possible to learn the feedback weights using the noise fluctuations of  $\mathbf{u}$ , while simultaneously letting the controller push the network towards the output target for learning the forward weights. (ii) The scaling factor  $(1 + \tau_v/\tau_\epsilon)$  enables the network to learn an optimal feedback connectivity pattern, while having a finite neural time constant  $\tau_v$ , which introduces delays in the network. As the correlation signal decreases for increasing time delays between the pre- and postsynaptic noise fluctuations, this needs to be compensated by an appropriate scaling factor.

**Feedback weights satisfy Conditions 1 and 2.** Theorem 3 shows that under simplifying conditions, the plasticity rule Eq. (9) drives the feedback connectivity  $Q$  to satisfy Conditions 1 and 2 (full proof in App. C).

**Theorem 3.** *Assume a separation of timescales  $\tau_v, \tau_\epsilon \ll \tau_u \ll \tau_f \ll \tau_Q$ , integral control ( $k_p = 0$ ), stable network dynamics, and small noise perturbations  $\sigma \ll \|\mathbf{v}_i\|$ . Then, for a fixed data sample, the feedback plasticity dynamics (9) let the first moment of  $Q$  converge approximately towards*

$$E[Q_{ss}] \approx J_{ss}^T M, \quad (10)$$

with  $M$  a positive definite symmetric matrix, which satisfies Conditions 1 and 2.

To ensure that the network is stable at all stages, we pre-train the feedback weights with a large  $\alpha$  to guarantee network stability while aligning the feedback weights with Eq. (10), after which we start the full network training with a small  $\alpha$ . Theorem 3 considers training the feedback weights on a single input sample. In reality, however, multiple input samples are used for training. For linear networks, the network Jacobian  $J_{ss}$  is independent of the input samples and Theorem 3 holds exactly. For nonlinear networks,  $J_{ss}$  will change depending on the input, hence, the feedback plasticity (9) will drive  $Q$  to align with some average of Eq. (10) for all input samples.

**Debiasing the forward weight updates.** While we need noise correlations for training the feedback weights, these can bias the forward weight updates (see App. C.4). We solve this challenge by low-pass filtering the pre-synaptic plasticity signal for the forward weight updates, such that

<sup>1</sup>We high-pass filter  $\mathbf{u}(t)$  by subtracting the exponentially averaged signal from it:  $\mathbf{u}_{\text{hp}}(t) = \mathbf{u}(t) - \mathbf{u}_{\text{lp}}(t)$ , with  $\mathbf{u}_{\text{lp}} = \frac{1}{\tau_f} \int_0^t \exp(-\frac{1}{\tau_f}(t-t')) \mathbf{u}(t') dt'$  and  $\tau_f$  the filtering time constant.

the pre-synaptic noise fluctuations are removed:

$$\tau_W \frac{d}{dt} W_i(t) = \left( \phi(\mathbf{v}_i(t)) - \phi(W_i \mathbf{r}_{i-1}(t)) \right) \mathbf{r}_{i-1}^T, \quad (11)$$

with  $\mathbf{r}_{i-1}(t) = \frac{1}{\tau_f} \int_0^t \exp(-\frac{1}{\tau_f}(t-t^\theta)) \mathbf{r}_i(t^\theta) dt^\theta$  an exponential moving average of  $\mathbf{r}_i$  with filtering time constant  $\tau_f$ .

## 5. Experiments

Here, we validate our theoretical findings with empirical results. By studying in detail a synthetic toy example and modest computer-vision benchmarks, we show that the minimizing control framework succeeds in minimizing the training loss  $\mathcal{L}$ , and that Strong-DFC equipped with single-phase learning succeeds in approximating the gradient of  $\mathcal{H}$ , thereby validating itself as a principled CA method. Finally, we show that Strong-DFC is significantly more noise-robust compared to standard DFC, due to its strong feedback influences on neural activity.<sup>2</sup>

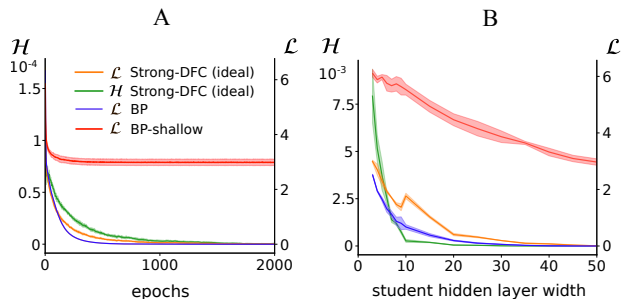
### Strong-DFC performs CA by minimizing feedback control.

We start with validating our main theoretical claim that minimizing control is a suitable new theoretical framework for learning. To this end, we first focus on a low-dimensional teacher-student nonlinear function approximation problem, which can be efficiently simulated.

In this problem, the goal of learning is to reduce the squared error  $\mathcal{L} = \mathbb{E}_{\mathbf{x} \sim \mathcal{N}(0, I)} [\|f_{\text{trgt}}(\mathbf{x}) - \mathbf{v}_{L, \text{ss}}(\mathbf{x})\|^2]$  between the target  $f_{\text{trgt}}(\mathbf{x})$  provided by the teacher network we wish to approximate and the steady-state last-layer activity of our student network, which is a function of the input  $\mathbf{x}$ . The student network is trained with Strong-DFC, and its input layer activity  $\mathbf{r}_0$  is clamped to the same random pattern  $\mathbf{x} \sim \mathcal{N}(0, I)$  used to probe the teacher function. To confirm that Theorem 1 and Proposition 2 hold in practice, we study an idealized setting. We disable neural noise ( $\sigma = 0$ ), simulate the neural dynamics (8) to equilibrium, and then apply a single weight change to  $W$  according to the steady-state Strong-DFC update (7). Moreover, to ensure that Condition 1 is satisfied in this idealized setting at all points during training, we manually set the feedback weights equal to the transposed network Jacobian:  $Q = J^T|_{\mathbf{v}=\mathbf{v}(t)}$  (see Condition 1). We use Euler’s forward method to simulate the dynamics and we initialize the network state to the feedforward prediction for each datapoint.

Our experiments confirm that the surrogate loss  $\mathcal{H}$  is reduced as training progresses, eventually reaching a low value (of the order of  $10^{-7}$ ), in accordance with Theorem 1 (see Fig. 2A). Crucially, we also observe that as  $\mathcal{H}$  decreases,

<sup>2</sup>Source code for all experiments is available at: [https://github.com/mariacer/strong\\_dfc](https://github.com/mariacer/strong_dfc).



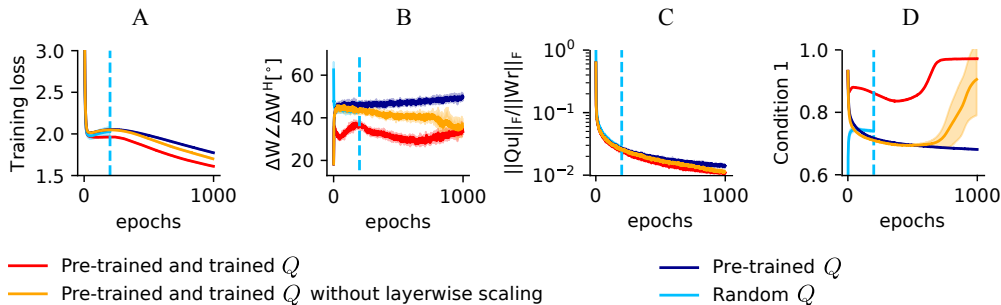
**Figure 2. Strong-DFC minimizes the original training loss and the amount of control.** (A) The training loss  $\mathcal{L}$  as well as the surrogate loss  $\mathcal{H}$  indicating the amount of control received are minimized during training with Strong-DFC and ideal feedback weights. The results are obtained in a nonlinear student-teacher setting with a fixed teacher of size 30-10-10-10-5, and an over-parameterized student of size 30-50-50-50-5, both equipped with tanh nonlinearities. (B) As hinted by Proposition 2, the amount of control  $\mathcal{H}$  at the end of training depends on how well  $\mathcal{L}$  can be minimized, and is thus dependent on the capacity of the network. Same setting as in (A), with the loss at convergence plotted in function of the width of the hidden layers of the student that varies from 3 to 50. We compare the Strong-DFC results against back-propagation (BP) and a student network of the same architecture but where only the last layer is trained (BP-shallow).

so does the actual loss of interest  $\mathcal{L}$ . At the end of training,  $\mathcal{L}$  is of the order  $10^{-3}$ , which supports our claim that the minima of the two loss functions coincide (cf. Proposition 2). Furthermore, by varying the width of the hidden layers of the student (Fig. 2B), we show that the amount of control  $\mathcal{H}$  required at the end of training depends on how well the original training loss  $\mathcal{L}$  can be minimized, with more over-parameterized networks reaching lower loss  $\mathcal{L}$  and requiring less help  $\mathcal{H}$ . We note that such low  $\mathcal{L}$  cannot be achieved by a linear function approximator, nor by a shallow student that only learns its output weights. This confirms that the ideal version of Strong-DFC performs useful credit assignment. We conclude that learning by minimizing control succeeds in solving our nonlinear function approximation problem.

We confirm the above results for the idealized setting on the MNIST dataset (LeCun et al., 2010) (Table 1). Here we also observe that both the original training loss  $\mathcal{L}$  and the surrogate loss  $\mathcal{H}$  reach low values, which in turn translates into high testing accuracy. Overall, these results confirm that the ideal version of Strong-DFC performs useful credit assignment in more challenging problems.

### Feedback learning dynamically tracks feedforward learning.

Having verified that Strong-DFC fulfills the minimizing control objective, and that this solves our task, we move to a more challenging setting where feedback weights are no longer manually fixed to ideal values, and where neural activity is noisy. The learning problem is otherwise



**Figure 3. Strong-DFC simulations for a nonlinear student-teacher regression task.** Standard Strong-DFC, where the feedback weights  $Q$  are both pre-trained before training the forward weights and trained simultaneously with the forward weights according to Eq. 9 (red), is compared to a set of controls where the feedback weights are: trained without the  $(1 + \tau_v/\tau_\epsilon)^{(L-i)}$  scaling term in the updates (orange), only pre-trained (dark blue), or kept fixed to a random initialization during the entire simulation (light blue). The blue vertical dotted lines indicate the moment where training with random feedback weights leads to instabilities. (A) Strong-DFC successfully reduces the loss  $L$  during training. (B) When simultaneously learning the feedback weights, the forward weight updates of Strong-DFC align well with the ideal updates following the gradient of  $H$ , showing that the amount of control is being minimized with our learning rules. (C) This is achieved through the use of *strong* feedback, which allows driving the network to the desired output even at early stages of training when the ratio of feedback ( $kQuk_2^2$ ) to forward ( $kWrk_2^2$ ) input is high. (D) The single-phase learning scheme succeeds in driving the feedback weights to satisfying Condition 1, enabling efficient training of the forward weights. We measure the ratio between the norm of  $Q$  projected into the row space of  $J_{ss}$ , and  $Q$  itself, with a value of 1.0 representing perfect compliance (see App. E). The results are obtained in a student-teacher setting with a fixed teacher of size 30-10-10-10-5, and an overparameterized student of size 30-50-50-50-5 with tanh nonlinearities. Shaded regions indicate standard deviations across five runs.

**Table 1.** Strong-DFC (with ideal feedback weights) successfully minimizes the original loss  $L$  and the surrogate loss  $H$  in MNIST. Results (mean  $\pm$  std, 5 seeds) using a multi-layer perceptron with three hidden layers of 256 neurons. Values are reported for the epoch with lowest training loss out of 40.

	Strong-DFC (ideal)		BP	
Train loss $\mathcal{L}$	8.09	$0.35 \cdot 10^{-2}$	3.38	$1.30 \cdot 10^{-5}$
Test loss $\mathcal{L}$	1.71	$0.10 \cdot 10^{-1}$	9.60	$0.04 \cdot 10^{-2}$
Train error	0.07	0.09%	0.00	0.00%
Test error	1.98	0.11%	1.78	0.04%
Train loss $\mathcal{H}$	2.64	$3.29 \cdot 10^{-5}$	N/A	

unchanged. In these experiments, we now adjust the feedback weights  $Q$  according to our anti-Hebbian plasticity rule (9) concurrently as the forward weights evolve through our always-on, debiased plasticity rule (11). For computational efficiency, we accumulate our weight changes over time and apply them once the neural steady-state is reached. We compare Strong-DFC to various ablations on the feedback weight learning (Fig. 3), and focus on the initial stage of learning as this provides the clearest differences. We observe that Strong-DFC successfully decreases the training loss (Fig. 3A) through updates that approximately align with the negative gradient of  $\mathcal{H}$  (Fig. 3B). Furthermore, our single-phase feedback learning rule succeeds in dynamically tracking the ongoing feedforward learning process and in finding values of  $Q$  which approximately respect Condition 1 (Fig. 3D). We note that throughout our sim-

ulation the feedforward and feedback inputs to any given neuron are of the same order of magnitude during early training, which confirms that the network operates in the strong feedback regime (Fig. 3C). Importantly, training performance, alignment with the gradient of  $\mathcal{H}$  and compliance with Condition 1 are all negatively impacted whenever the feedback weights  $Q$  are not trained during the simulation. Furthermore, without any pre-training, updates quickly lead to numerical instabilities. This indicates that purely random feedback mappings are not sufficient for activity-dependent learning with strong feedback modulation, in contrast to random feedback mappings that are solely used for plasticity (Lillicrap et al., 2016; Nøkland, 2016).

### Strong-DFC exhibits robustness to noise and competitive performance in standard computer vision benchmarks.

We now investigate if Strong-DFC can successfully perform credit assignment on a deeper neural network model, that we apply to two standard image classification benchmarks, MNIST and Fashion-MNIST (Xiao et al., 2017). Moreover, we set the focus of our investigation on noisy neural states, where  $\sigma$  in (8) is non-negligible ( $\sigma \sim 10^{-2}$ ), to determine whether the strong feedback introduced in Strong-DFC enables learning in the high-noise regime, and whether our plasticity rule (9) can use the fluctuations to learn feedback connections that are useful for credit assignment on deeper networks. In this section, we study a fully-connected three-hidden-layer neural network (256 units per layer) with a final softmax classification layer and choose the cross-entropy as our loss function  $\mathcal{L}$ , which can be handled in

our framework as detailed in Appendix A.6. When using Strong-DFC, we learn feedback and feedforward weights concurrently with always-on plasticity rules; plasticity is only interrupted when switching patterns and re-initializing the network states. To ensure network stability at the beginning of training, we pre-train the feedback weights using Eq. (9) and a high controller leak rate  $\alpha$ .

Table 2. Computer vision results. Test errors (mean  $\pm$  std, 5 seeds) obtained after 40 training epochs on a multi-layer perceptron with three hidden layers of 256 neurons each. For DFC, results are provided without noise (original) and with noise during the forward weight training. For Strong-DFC, we consider the (original) single-phase setting with noise and a two-phase setting without noise in the forward weights training phase.

	MNIST-		standard (%) fashion (%)	
BP	1.83	0.11	10.60	0.44
Strong-DFC (with noise)	2.19	0.05	12.07	0.16
DFC (with noise)	15.15	0.44	16.29	0.41
Strong-DFC (no noise)	1.98	0.05	11.36	0.17
DFC (no noise)	2.09	0.10	11.31	0.14

In Table 2, we compare the test set prediction error rates of Strong-DFC to DFC, as well as to standard backpropagation (BP), which we provide as a reference baseline. We optimize the hyperparameters of each method independently, for best performance on a validation set of 5000 datapoints. We use BP to train a noiseless network, and DFC to train both a noiseless and a noisy network, where we add noise to the dynamics according to Eq. (8). In the noisy setting, we equip DFC with the debiased plasticity rule (11) for fair comparison. As previously reported (Meulemans et al., 2021), in the noiseless case, the test-set performance of DFC is close to that achieved by BP. However, the situation changes in the large- $\sigma$  regime: the performance of (weak feedback) DFC drastically drops, whereas Strong-DFC is undisturbed on MNIST and not severely impaired on Fashion-MNIST (we further illustrate this in a student-teacher regression setting in Appendix E.2). When we train the forward weights with Strong-DFC in a separate phase without noise, its performance is competitive to both DFC and BP. Moreover, we observed that without feedback plasticity, we did not succeed in training our network when starting from a random initial  $Q$ , as the network was unstable. Our feedback plasticity thus succeeds in pushing the initial network state to an appropriate one and in keeping it throughout learning.

## 6. Discussion

Modern deep neural networks are invariably learned by backpropagation-of-error. While backpropagation has proven to be a highly effective gradient-based credit as-

ignment method for training deep artificial neural networks, it is unsatisfactory as a model of learning in cortical networks in at least three important aspects. First, it needs precisely-symmetric forward and feedback connections. Second, backpropagated errors do not influence neural activity. Third, it requires two precisely-clocked phases. Instead of attempting to remedy these issues one by one, we propose a fundamentally new approach for deep learning, that overcomes all of these by design, and provide first results indicating that it is an effective method for credit assignment.

Our study shows that these three critical issues can be overcome at once within the framework of gradient-based optimization, by casting learning as a control minimization problem. For this, we first augment standard deep neural networks with a feedback controller, which dynamically changes the hidden neural activity to reach the desired output. Then, we derive a synaptic plasticity rule with the objective of minimizing the influence of the feedback controller on the neural activity. This principled derivation led us to Strong-DFC, a variation of the recently proposed DFC method which inherits all of its appealing properties as a biologically plausible deep learning framework. Like DFC, our weight update rule is local in space and time for a broad range of (not necessarily symmetric) feedback connectivity patterns and neural feedback controller architectures, and it can be interpreted as a biologically inspired dendritic voltage-dependent synaptic plasticity rule (Urbanczik & Senn, 2014). Importantly, however, top-down feedback signals are no longer required to be vanishingly small for Strong-DFC.

This strong feedback influence sets our minimizing control framework apart from previous theories of how the brain might estimate objective function gradients by measuring small changes in neural activity, generated by ‘nudging’ the network towards a slightly better state (Hinton & McClelland, 1988; O’Reilly, 1996; Xie & Seung, 2003; Scellier & Bengio, 2017; Sacramento et al., 2018; Lillicrap et al., 2020; Meulemans et al., 2021). In biology, such theories might face a fundamental obstacle during learning, since cortical activity is notoriously noisy (Rusakov et al., 2020), making it hard to estimate small activity changes as plasticity signals. Moreover, they are at odds with a growing number of experimental neuroscience observations reporting large neural activity changes in the presence of unexpected or novel events, which are thought to drive synaptic plasticity (Keller & Mrcic-Flogel, 2018), and the existence of top-down feedback signals that are sufficient to drive secondary receptive field responses in pyramidal neurons (Keller et al., 2020). Our minimizing control framework shows how to perform principled mathematical optimization of an objective function, while being consistent with these findings.



A long-standing question in neuroscience is whether noise plays a computational role in the brain – whether it is ‘a bug or a feature’ of biological neural circuits (Rusakov et al., 2020). Our study suggests that noise could play a key role in ensuring that gradient-based credit assignment information is transmitted backwards through the network. Building on the noise-tolerance of Strong-DFC, we leverage noise to concurrently and seamlessly learn both feedforward and feedback weights so that Condition 1 is constantly fulfilled. In contrast to standard two-phase learning of forward and backward connections (Akrouf et al., 2019; Lee et al., 2015; Meulemans et al., 2021), our plasticity model does not require any pauses, distinct phases, plasticity switches, or alternations between noise- and input-driven modes. Our always-on plasticity rules are arguably more biologically plausible than alternative two-phase learning algorithms, and simpler to implement in neuromorphic hardware.

In our pyramidal cell model, top-down apical synapses transmit feedback information which instructs the plasticity of bottom-up basal synapses, an idea that can be traced back to the seminal work of Körding & König (2001). Interpreted as such, our results suggest that the plasticity rules for top-down and bottom-up connections might display major differences, which is experimentally testable. Both rules include a voltage-dependent postsynaptic factor, in agreement with phenomenological plasticity models (Clopath & Gerstner, 2010). However, the presynaptic components of our apical and basal plasticity rules are high- and low-pass filtered, respectively. Such presynaptic activity filters are a recurring element in reward-driven synaptic plasticity rules (Seung, 2003). Moreover, our apical plasticity rule is anti-Hebbian and its postsynaptic term depends only on the local dendritic voltage, not on the somatic spiking activity. This may be consistent with the fact that, for pyramidal cells, backpropagating action potentials often fail to reach the apical dendrite, questioning their role as the main drivers of apical synaptic plasticity (Spruston, 2008; Gambino et al., 2014).

We highlight that, despite the close methodological connections between DFC and Strong-DFC, our minimizing control framework is fundamentally different from the underlying theory of DFC and many other recent biologically plausible learning methods that link their weight updates to a new variant of Gauss-Newton optimization with a mini-batch size of one (Meulemans et al., 2021; 2020; Podlaski & Machens, 2020; Bengio, 2020). While the specific optimization characteristics of these methods are not yet fully understood, our minimizing control framework can be cast as gradient descent on a surrogate loss function. Hence, the rich field studying the application of stochastic gradient descent to neural networks (Bottou et al., 2018) can be applied on our new loss to uncover the optimization characteristics of Strong-DFC. Furthermore, our minimizing control ap-

proach draws from control theory, but it is distinct from standard optimal control. In particular, it is well known that the adjoint state method from optimal control is equivalent to backpropagation on the supervised loss function, when applied to deep feedforward neural networks (LeCun et al., 1988), and hence different from our minimizing control framework that performs gradient descent on a surrogate loss function.

In practice, the Strong-DFC updates do not exactly follow the negative gradient of  $\mathcal{H}$ , due to (i) the limited number of iterations for training the feedback weights, (ii) the limited capacity of the linear pathway  $Q$  to satisfy Condition 1 for every input sample, and (iii) the difference between the steady-state weight update (7) and the Strong-DFC weight update (4). Despite these challenges, Fig 3 shows that Strong-DFC approximates the theory well. Future work can further improve Strong-DFC by investigating alternative feedback mappings and network architectures. As Strong-DFC uses continuous dynamics, it is costly to simulate on conventional deep learning hardware, preventing it from being tested on large-scale machine learning benchmarks. A promising solution is to implement Strong-DFC on analog and neuromorphic hardware, where the dynamics can coincide with the physical dynamics of the analog components. Consequently, this can make Strong-DFC an attractive principled approach for CA on analog deep learning implementations, commonly used in Edge-AI and other low-energy applications of deep learning (Xiao et al., 2020; Misra & Saha, 2010).

By providing a novel optimization approach for deep neural networks that differs from the usual direct minimization of the output loss, our minimizing control framework allowed us to derive a novel gradient-based credit assignment method that relies on strong feedback influence. Instantiating our framework with Strong-DFC naturally leads to biologically desired characteristics, such as single-phase learning, flexible feedback connectivity requirements, and local learning rules, further underlining its promise for understanding credit assignment in the brain.

## Acknowledgements

This work was supported by the Swiss National Science Foundation (B.F.G. CRSII5-173721 and 315230\_189251), ETH project funding (B.F.G. ETH-20 19-01) and the Human Frontiers Science Program (RGY0072/2019). João Sacramento was supported by an Ambizione grant (PZ00P3\_186027) from the Swiss National Science Foundation. We would like to thank Nicolas Zucchet for his valuable feedback and engaging discussions on the optimization theory aspects of the minimizing control framework; and Aditya Gilra for insightful discussions on the learning through control framework.

**References**

- Akroud, M., Wilson, C., Humphreys, P., Lillicrap, T., and Tweed, D. B. Deep learning without weight transport. In *Advances in Neural Information Processing Systems 32*, pp. 974–982, 2019.
- Alemi, A., Machens, C., Denève, S., and Slotine, J.-J. Learning arbitrary dynamics in efficient, balanced spiking networks using local plasticity rules. *AAAI Conference on Artificial Intelligence (AAAI)*, 2018.
- Belkin, M., Hsu, D., Ma, S., and Mandal, S. Reconciling modern machine-learning practice and the classical bias–variance trade-off. *Proceedings of the National Academy of Sciences*, 116(32):15849–15854, 2019.
- Bengio, Y. Deriving differential target propagation from iterating approximate inverses. *arXiv preprint arXiv:2007.15139*, 2020.
- Bottou, L., Curtis, F. E., and Nocedal, J. Optimization methods for large-scale machine learning. *SIAM Review*, 60(2):223–311, 2018.
- Bourdoukan, R. and Deneve, S. Enforcing balance allows local supervised learning in spiking recurrent networks. *Advances in Neural Information Processing Systems*, 28: 982–990, 2015.
- Clopath, C. and Gerstner, W. Voltage and spike timing interact in stdp—a unified model. *Frontiers in synaptic neuroscience*, 2:25, 2010.
- Crick, F. The recent excitement about neural networks. *Nature*, 337(6203):129–132, 1989.
- Denève, S., Alemi, A., and Bourdoukan, R. The brain as an efficient and robust adaptive learner. *Neuron*, 94(5): 969–977, 2017.
- Gambino, F., Kehayas, V., Baptista, D., Tatti, R., Carleton, A., and Holtmaat, A. Sensory-evoked LTP driven by dendritic plateau potentials in vivo. *Nature*, 515(7525): 116–119, 2014.
- Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press, 2014.
- Gilra, A. and Gerstner, W. Predicting non-linear dynamics by stable local learning in a recurrent spiking neural network. *Elife*, 6:e28295, 2017.
- Hinton, G. E. and McClelland, J. L. Learning representations by recirculation. In *Advances in Neural Information Processing Systems*, pp. 358–366, 1988.
- Jordan, R. and Keller, G. B. Opposing influence of top-down and bottom-up input on excitatory layer 2/3 neurons in mouse primary visual cortex. *Neuron*, 108(6):1194–1206, 2020.
- Keller, A. J., Roth, M. M., and Scanziani, M. Feedback generates a second receptive field in neurons of the visual cortex. *Nature*, 582(7813):545–549, 2020.
- Keller, G. B. and Mrsic-Flogel, T. D. Predictive processing: a canonical cortical computation. *Neuron*, 100(2):424–435, 2018.
- Körding, K. P. and König, P. Supervised and unsupervised learning with two sites of synaptic integration. *Journal of Computational Neuroscience*, 11(3):207–215, 2001.
- Lansdell, B. J., Prakash, P., and Kording, K. P. Learning to solve the credit assignment problem. In *International Conference on Learning Representations*, 2020.
- LeCun, Y., Touresky, D., Hinton, G., and Sejnowski, T. A theoretical framework for back-propagation. In *Proceedings of the 1988 Connectionist Models Summer School*, volume 1, pp. 21–28. Morgan Kaufmann, 1988.
- LeCun, Y., Cortes, C., and Burges, C. MNIST handwritten digit database. *AT&T Labs*. Available: <http://yann.lecun.com/exdb/mnist>, 2010.
- LeCun, Y., Bengio, Y., and Hinton, G. Deep learning. *Nature*, 521(7553):436, 2015.
- Lee, D.-H., Zhang, S., Fischer, A., and Bengio, Y. Difference target propagation. In *Joint european conference on machine learning and knowledge discovery in databases*, pp. 498–515. Springer, 2015.
- Lillicrap, T. P., Cownden, D., Tweed, D. B., and Akerman, C. J. Random synaptic feedback weights support error backpropagation for deep learning. *Nature Communications*, 7:13276, 2016.
- Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., and Hinton, G. Backpropagation and the brain. *Nature Reviews Neuroscience*, pp. 1–12, 2020.
- Lu, T.-T. and Shiou, S.-H. Inverses of  $2 \times 2$  block matrices. *Computers & Mathematics with Applications*, 43(1-2): 119–129, 2002.
- Meulemans, A., Carzaniga, F., Suykens, J., Sacramento, J. a., and Grewe, B. F. A theoretical framework for target propagation. *Advances in Neural Information Processing Systems*, 33:20024–20036, 2020.
- Meulemans, A., Farinha, M. T., Ordóñez, J. G., Aceituno, P. V., Sacramento, J., and Grewe, B. F. Credit assignment in neural networks through deep feedback control. *arXiv preprint arXiv:2106.07887*, 2021.

- Misra, J. and Saha, I. Artificial neural networks in hardware: A survey of two decades of progress. *Neurocomputing*, 74(1-3):239–255, 2010.
- Nøkland, A. Direct feedback alignment provides learning in deep neural networks. In *Advances in neural information processing systems*, pp. 1037–1045, 2016.
- O’Reilly, R. C. Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, 8(5):895–938, 1996.
- Payeur, A., Guerguiev, J., Zenke, F., Richards, B., and Naud, R. Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. *Nature neuroscience*, 24(5):1546, 2021.
- Podlaski, W. F. and Machens, C. K. Biological credit assignment through dynamic inversion of feedforward networks. *Advances in Neural Information Processing Systems 33*, 2020.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. Learning representations by back-propagating errors. *Nature*, 323(6088):533, 1986.
- Rusakov, D. A., Savtchenko, L. P., and Latham, P. E. Noisy synaptic conductance: bug or a feature? *Trends in Neurosciences*, 43(6):363–372, 2020.
- Sacramento, J., Costa, R. P., Bengio, Y., and Senn, W. Dendritic cortical microcircuits approximate the backpropagation algorithm. In *Advances in Neural Information Processing Systems 31*, pp. 8721–8732, 2018.
- Särkkä, S. and Solin, A. *Applied stochastic differential equations*, volume 10. Cambridge University Press, 2019.
- Scellier, B. and Bengio, Y. Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in computational neuroscience*, 11:24, 2017.
- Seung, H. S. Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron*, 40(6):1063–1073, 2003.
- Spruston, N. Pyramidal neurons: dendritic structure and synaptic integration. *Nature Reviews Neuroscience*, 9(3):206–221, 2008.
- Urbanczik, R. and Senn, W. Learning by the dendritic prediction of somatic spiking. *Neuron*, 81(3):521–528, 2014.
- Werbos, P. J. Applications of advances in nonlinear sensitivity analysis. In *System modeling and optimization*, pp. 762–770. Springer, 1982.
- Whittington, J. C. and Bogacz, R. An approximation of the error backpropagation algorithm in a predictive coding network with local hebbian synaptic plasticity. *Neural computation*, 29(5):1229–1262, 2017.
- Xiao, H., Rasul, K., and Vollgraf, R. Fashion-MNIST: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*, 2017.
- Xiao, T. P., Bennett, C. H., Feinberg, B., Agarwal, S., and Marinella, M. J. Analog architectures for neural network acceleration based on non-volatile memory. *Applied Physics Reviews*, 7(3):031301, 2020.
- Xie, X. and Seung, H. S. Equivalence of backpropagation and contrastive Hebbian learning in a layered network. *Neural Computation*, 15(2):441–454, 2003.
- Yamins, D. L. and DiCarlo, J. J. Using goal-driven deep learning models to understand sensory cortex. *Nature neuroscience*, 19(3):356–365, 2016.

## A. Proofs and discussion for Section 3

### A.1. Proof Theorem 1

In this appendix section, we prove Theorem 1, which we copy below for convenience.

**Theorem S1.** *Assuming  $J_{ss}Q + \tilde{\alpha}I$  is invertible, the total derivative of  $\mathcal{H}$  w.r.t.  $W_i$  is given by:*

$$\frac{d\mathcal{H}}{dW_i} = - \sum_b J_{i,ss}^{(b)T} \left( J_{ss}^{(b)}Q + \tilde{\alpha}I \right)^{-1} Q^T Q \mathbf{u}_{ss}^{(b)} \mathbf{r}_{i-1,ss}^{(b)T}, \quad (12)$$

with  $J_{i,ss}$ ,  $\frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_i}$  and  $J_{ss}$ ,  $\left[ \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1}, \dots, \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L} \right]$ , both evaluated at steady state, and  $\tilde{\alpha}$ ,  $\frac{\alpha}{1+k\alpha}$ .

Furthermore, assuming Condition 1 holds, the network is stable,  $\alpha \rightarrow 0$ , and  $J_{ss}Q + \tilde{\alpha}I$  is invertible, then the following steady-state (ss) updates for the forward weights, averaged over the batch samples,

$$\Delta W_{i,ss} = \eta (\mathbf{v}_{i,ss} - \mathbf{v}_{i,ss}^{\text{ff}}) \mathbf{r}_{i-1,ss}^T, \quad (13)$$

with  $\eta$  the stepsize, align with the negative gradient  $-\frac{d\mathcal{H}}{dW_i}$ .

**Condition S1.** The column space of  $Q$  is equal to the row space of  $J_{ss}$ ,  $\left[ \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1}, \dots, \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L} \right]$  at steady state.

The intuition behind the proof is as follows. First, we compute in Lemma S2 the gradient of  $\mathcal{H}$  using the Implicit Function Theorem, as this provides us with the means to calculate how the control signal  $\mathbf{u}_{ss}$  at steady state changes when  $W_i$  is changed. In Lemma S3, we show that  $\frac{d\mathcal{H}}{dW_i}$  collapses to a local update if Condition 1 is satisfied. Finally, we prove Theorem 1 by showing that this local update is equal to the steady-state weight update in Eq. (7).

**Lemma S2.** *The gradient of  $\mathcal{H}$  w.r.t.  $\theta$  for a single sample, evaluated at steady state, is given by*

$$\frac{1}{2} \frac{d\|Q\mathbf{u}_{ss}\|_2^2}{d\theta} = -\mathbf{u}_{ss}^T Q^T Q (J_{ss}Q + \tilde{\alpha}I)^{-1} J_{ss} R_{ss}^T, \quad (14)$$

with  $\theta = [\vec{W}_1^T, \dots, \vec{W}_L^T]^T$  and  $R_{ss}$  defined below.

$$R_{ss}^T = \begin{bmatrix} \mathbf{r}_0^T \otimes I & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{r}_{L-1}^T \otimes I \end{bmatrix} \quad (15)$$

with  $\otimes$  the Kronecker product.

*Proof.* Using the chain rule, we get the following expression for the gradient of interest:

$$\frac{1}{2} \frac{d\|Q\mathbf{u}_{ss}\|_2^2}{d\theta} = \mathbf{u}_{ss}^T Q^T Q \frac{d\mathbf{u}_{ss}}{d\theta} \quad (16)$$

Now we proceed to find  $\frac{d\mathbf{u}_{ss}}{d\theta}$ . At steady state, the network and controller dynamics result in the following equilibrium equations:

$$\mathbf{v}_{i,ss} = W_i \phi(\mathbf{v}_{i-1,ss}) + Q_i \mathbf{u}_{ss}, \quad 1 \leq i \leq L \quad (17)$$

$$\mathbf{u}_{ss} = k \mathbf{e}_{ss} + \mathbf{u}_{ss}^{\text{int}} \quad (18)$$

$$0 = \mathbf{e}_{ss} - \alpha \mathbf{u}_{ss}^{\text{int}} \quad (19)$$

$$\mathbf{e}_{ss} = \mathbf{r}_L - \mathbf{r}_{L,ss} \quad (20)$$

The equilibrium solutions in Eq. (17) can be used to obtain an explicit function of  $\mathbf{r}_L$  w.r.t.  $\theta$  and  $\mathbf{u}_{ss}$ , which in turn can be used to obtain an explicit function of  $\mathbf{e}_{ss}$  i.f.o.  $\theta$  and  $\mathbf{u}_{ss}$ . Combining all equations results in the following implicit function relating  $\mathbf{u}_{ss}$  to  $\theta$ :

$$F(\mathbf{u}_{ss}, \theta), \quad \mathbf{e}_{ss}(\mathbf{u}_{ss}, \theta) - \tilde{\alpha} \mathbf{u}_{ss} = 0, \quad (21)$$



with  $\tilde{\alpha} = \frac{\alpha}{1+\alpha k}$ . If  $\frac{\partial F}{\partial \mathbf{u}_{ss}}$  is invertible, we can use the Implicit Function Theorem to obtain the total derivative of  $\mathbf{u}_{ss}$  w.r.t.  $\boldsymbol{\theta}$ :

$$\frac{d\mathbf{u}_{ss}}{d\boldsymbol{\theta}} = - \left( \frac{\partial F}{\partial \mathbf{u}_{ss}} \right)^{-1} \frac{\partial F}{\partial \boldsymbol{\theta}}. \quad (22)$$

Investigating each part separately, we get

$$\frac{\partial F}{\partial \mathbf{u}_{ss}} = - \frac{\partial \mathbf{r}_L}{\partial \mathbf{u}} - \tilde{\alpha} I \quad (23)$$

$$= - \sum_{i=1}^L \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_i} \frac{\partial \mathbf{v}_i}{\partial \mathbf{u}} - \tilde{\alpha} I \quad (24)$$

$$= -J_{ss} Q - \tilde{\alpha} I, \quad (25)$$

with all derivatives evaluated at the steady-state activations, and

$$\frac{\partial F}{\partial \boldsymbol{\theta}} = \frac{\partial \mathbf{r}_L}{\partial \boldsymbol{\theta}} \quad (26)$$

$$= -J_{ss} R_{ss}^T, \quad (27)$$

with  $R_{ss}^T$  defined in Eq. (15). Bringing everything together results in

$$\frac{d\mathbf{u}_{ss}}{d\boldsymbol{\theta}} = - (J_{ss} Q + \tilde{\alpha} I)^{-1} J_{ss} R_{ss}^T \quad (28)$$

$$\frac{1}{2} \frac{d\|Q\mathbf{u}_{ss}\|_2^2}{d\boldsymbol{\theta}} = -\mathbf{u}_{ss}^T Q^T Q (J_{ss} Q + \tilde{\alpha} I)^{-1} J_{ss} R_{ss}^T. \quad (29)$$

□

**Lemma S3.** When Condition 1 is satisfied and  $J_{ss} Q$  is invertible,  $\frac{d\|Q\mathbf{u}_{ss}\|_2^2}{dW_i}$  results in a local update rule in the limit of  $\tilde{\alpha} \rightarrow 0$ :

$$\lim_{\tilde{\alpha} \rightarrow 0} \frac{1}{2} \frac{d\|Q\mathbf{u}_{ss}\|_2^2}{dW_i} = -Q_i \mathbf{u}_{ss} \mathbf{r}_i^T. \quad (30)$$

*Proof.* Using a similar proof as Lemma S2 in Meulemans et al. (2021), we have that

$$\lim_{\tilde{\alpha} \rightarrow 0} (J_{ss} Q + \tilde{\alpha} I)^{-1} J_{ss} = Q^\gamma \quad (31)$$

iff  $\text{Row}(J_{ss}) = \text{Col}(Q)$ , with  $Q^\gamma$  the Moore-Penrose pseudoinverse of  $Q$ . Using the fact that  $Q^T Q Q^\gamma = Q^T$ , we have that

$$\lim_{\tilde{\alpha} \rightarrow 0} \frac{1}{2} \frac{d\|Q\mathbf{u}_{ss}\|_2^2}{d\boldsymbol{\theta}} = -\mathbf{u}_{ss}^T Q^T R_{ss}^T \quad (32)$$

iff  $\text{Row}(J_{ss}) = \text{Col}(Q)$  and  $J_{ss}$  is of full row rank. Going from the vectorized notation for  $\boldsymbol{\theta}$  towards the matrix notation for  $W_i$  gives us

$$\lim_{\tilde{\alpha} \rightarrow 0} \frac{1}{2} \frac{d\|Q\mathbf{u}_{ss}\|_2^2}{dW_i^T} = -Q \mathbf{u}_{ss} \mathbf{r}_i^T. \quad (33)$$

□

Now, we are ready to prove Theorem 1.

*Proof.* From the linearity of the derivative operator, we have that

$$\frac{d\mathcal{H}}{dW_i^T} = \sum_b \frac{1}{2} \frac{d\|Q\mathbf{u}_{ss}^{(b)}\|_2^2}{dW_i}. \quad (34)$$

Rewriting Lemma S2 from its vectorized form towards matrix form, and filling it in the above equation, proves the first part of Theorem 1:

$$\frac{d\mathcal{H}}{dW_i^T} = - \sum_b J_{i,ss}^{(b)T} \left( J_{ss}^{(b)} Q + \tilde{\alpha} I \right)^T Q^T Q \mathbf{u}_{ss}^{(b)} \mathbf{r}_{i,1,ss}^{(b)T}, \quad (35)$$

For the second part of the Theorem, the conditions of Lemma S3 are satisfied by the stated assumptions of this theorem. Hence, we have that:

$$\lim_{\tilde{\alpha} \downarrow 0} \frac{d\mathcal{H}}{dW_i^T} = - \sum_b Q_i \mathbf{u}_{ss}^{(b)} \mathbf{r}_{i,1,ss}^{(b)T}. \quad (36)$$

Finally, from the steady-state solution for the network dynamics, we have that  $\mathbf{v}_{i,ss} - \mathbf{v}_{i,ss}^{\text{fb}} = Q_i \mathbf{u}_{ss}$ , thereby concluding the proof.  $\square$

## A.2. Interpreting Condition 1

Here, we investigate in more depth why Condition 1 is needed in Theorem 1. The main line of argumentation goes as follows: if Condition 1 is not satisfied, the steady-state weight update given by Eq. (7) has components inside the nullspace of the network, and hence, can only partially influence the network output. The gradient of  $\|Q \mathbf{u}_{ss}\|_2^2$  is its steepest ascent direction, thus all the components of this gradient should influence the network output and consequently, the amount of needed control feedback, because otherwise there would exist a 'steeper' direction. Therefore, if the weight update given by Eq. (7) contains nullspace components, it cannot be fully aligned with the gradient.

Now, let us specify the *nullspace* of the network. As the dimension of the concatenated forward weights  $\boldsymbol{\theta}$  is higher than the dimension of the network output, there exist certain weight updates  $\Delta\boldsymbol{\theta}$  that do not result in any change to the output  $\mathbf{r}_L$  of the network, as they lie inside the nullspace of the network. For small weight updates, the nullspace of the network at steady state is connected to the network Jacobian, as we have that  $\Delta\mathbf{r}_L \approx J_{ss} R_{ss}^T \Delta\boldsymbol{\theta}$ , with  $R_{ss}$  as defined in Lemma S2. From this formulation, one can see that the nullspace of the network is equal to the nullspace of  $J_{ss} R_{ss}^T$ .

When vectorizing the steady-state weight updates (Eq. 7), we see that they lie inside the rowspace of  $Q^T R_{ss}^T$ . If we have that the column space of  $Q$  (hence, the rowspace of  $Q^T$ ) is equal to the row space of  $J_{ss}$ , the steady-state weight updates will always lie inside the rowspace of  $J_{ss} R_{ss}^T$ , and hence, can never lie inside the nullspace of the network.

The negative gradient of  $\|Q \mathbf{u}_{ss}\|_2^2$  is its steepest descent direction, thus all the components of the gradient should influence the network output and consequently the amount of needed control input. Therefore, if the weight update Eq. (7) contains nullspace components, it cannot be fully aligned with the gradient. Consequently, Condition 1 is a necessary condition for the steady-state update (Eq. 7) to align with the negative gradient. The sufficiency of Condition 1 follows from the proof of Theorem 1.

To illustrate the above arguments further, let us investigate in more detail what happens when we update the forward network parameters  $\boldsymbol{\theta}$ . Assuming a small weight update  $\Delta\boldsymbol{\theta}$ , we can do a first-order Taylor expansion of the steady-state network output after the weight update:

$$\mathbf{r}_{L,ss}^{(m+1)} \approx \mathbf{r}_{L,ss}^{(m)} + J_{\boldsymbol{\theta}} \Delta\boldsymbol{\theta} + J_{ss} Q \Delta \mathbf{u}_{ss}, \quad (37)$$

with  $J_{\boldsymbol{\theta}} = \frac{\partial \mathbf{r}_L}{\partial \boldsymbol{\theta}} = J_{ss} R_{ss}^T$  and  $m$  the training iteration. The weight update changes the amount of control needed to reach the output target, which we capture by introducing  $\Delta \mathbf{u}_{ss}$ . As the feedback controller drives the network to the supervised label during each training iteration (assuming  $\alpha \rightarrow 0$ ), we have that  $\mathbf{r}_{L,ss}^{(m+1)} = \mathbf{r}_{L,ss}^{(m)} = \mathbf{r}_L^{\text{true}}$  and hence

$$J_{ss} R_{ss}^T \Delta\boldsymbol{\theta} \approx -J_{ss} Q \Delta \mathbf{u}_{ss}. \quad (38)$$

The above equation relates the update  $\Delta\boldsymbol{\theta}$  to the change in steady-state control input  $Q \Delta \mathbf{u}_{ss}$ . In order to maximally reduce  $\mathcal{H}$  and hence  $\|Q \mathbf{u}_{ss}\|_2^2$ , the magnitude of the LHS should be as big as possible. As a result, we want  $\Delta\boldsymbol{\theta}$  to lie fully inside the rowspace of  $J_{ss} R_{ss}^T$ , because otherwise,  $\Delta\boldsymbol{\theta}$  will partially be lost in its nullspace, reducing the magnitude of the LHS. As  $\Delta\boldsymbol{\theta} = R_{ss} Q \Delta \mathbf{u}_{ss}$ , we have that the columnspace of  $R_{ss} Q$  needs to lie inside the rowspace of  $J_{ss} R_{ss}^T$ . This condition is equivalent to the columnspace of  $Q$  lying inside the rowspace of  $J_{ss}$ . Note that the above arguments should be interpreted on an intuitive level, a detailed mathematical consideration of the above arguments is more nuanced, but leads to the same high-level intuition.

### A.3. Effect of the nonlinearity $\phi$ on the weight updates

The main effect of including the nonlinearity  $\phi$  into the Strong-DFC weight updates (Eq. 4), in contrast to the steady-state updates (Eq. 7) of Theorem 1, is ensuring that saturated neurons refrain from updating their synaptic weights for the current data sample, as  $\phi(\mathbf{v}_i) - \phi(\mathbf{v}_i^{\text{ff}}) \approx 0$  in this case.

Meulemans et al. (2021) showed that including  $\phi$  in the learning rule is a useful heuristic for improving the alignment of the DFC weight updates with the theoretically ideal update (which in our case would be the negative gradient of  $\mathcal{H}$ ), when the column space condition on  $Q$  is not perfectly satisfied. Consequently, including  $\phi$  also improves the performance in this case. When the column space condition on  $Q$  is perfectly satisfied, including  $\phi$  in the learning rule makes no difference.

### A.4. Steady-state weight updates vs. continuous weight updates

Our theoretical results assume that the weight update is performed at steady state, giving rise to the update defined in Eq. (7). However, in practice, plasticity is always on, giving rise to the continuous plasticity dynamics defined in Eq. (4). Following Meulemans et al. (2021), we show that when the steady state is quickly reached, the steady-state update (7) approximates the resulting continuous update (4) well.

Assuming a separation of time scales between the network dynamics and the synaptic plasticity ( $\tau_v, \tau_u \ll \tau_W$ ), the accumulated weight update resulting from the plasticity dynamics (4) is given by

$$\int_0^{t_1} dW_i = \int_0^{t_{\text{ss}}} dW_i + \frac{t_1 - t_{\text{ss}}}{\tau_W} (\phi(\mathbf{v}_{i,\text{ss}}) - \phi(\mathbf{v}_{i,\text{ss}}^{\text{ff}})) \mathbf{r}_i^T \mathbf{1}_{,\text{ss}}. \quad (39)$$

Now, if the dynamics quickly settle towards their steady state, relative to the time interval the input is presented, i.e.,  $t_{\text{ss}} \ll t_1$ , the weight update is dominated by the second term in the RHS of the above equation, which corresponds to the steady-state weight update with  $\phi$  included. Furthermore, the controller quickly pushes the activity close to its steady state, with possibly some oscillations around it, so the first term  $\int_0^{t_{\text{ss}}} dW_i$  will approximate the steady state well.

In conclusion, the Strong-DFC weight update resulting from the plasticity dynamics described in Eq. (4) will approximate the steady-state weight update  $(\phi(\mathbf{v}_{i,\text{ss}}) - \phi(\mathbf{v}_{i,\text{ss}}^{\text{ff}})) \mathbf{r}_i^T \mathbf{1}_{,\text{ss}}$ . Combined with the arguments of Appendix A.3, we see that the Strong-DFC weight update approximates the steady-state update of Theorem 1 well.

### A.5. The minimizing control framework is compatible with various controller types

In Section 3, we introduced our theory on the minimizing control framework tailored towards a proportional intergral (PI) controller used by DFC. However, our theory can easily be extended to other controller types that reach the correct steady state.

In the proof of Theorem 1, the only assumption on the controller is its steady state equation (21), given by Eq. (21):

$$F(\mathbf{u}_{\text{ss}}, \boldsymbol{\theta}), \mathbf{e}_{\text{ss}}(\mathbf{u}_{\text{ss}}, \boldsymbol{\theta}) - \tilde{\alpha} \mathbf{u}_{\text{ss}} = 0.$$

Therefore, Theorem 1 applies to all stable controllers that give rise to this same steady-state equation,  $F(\mathbf{u}_{\text{ss}}, \boldsymbol{\theta})$ .

For example, pure proportional control,  $\mathbf{u}(t) = k\mathbf{e}(t)$  gives rise to the same steady-state equation  $F(\mathbf{u}_{\text{ss}}, \boldsymbol{\theta})$ , with  $\tilde{\alpha} = \frac{1}{k}$ . For pure integral control, we have the same steady-state equation  $F(\mathbf{u}_{\text{ss}}, \boldsymbol{\theta})$  but with  $\tilde{\alpha} = \alpha$ . If we add derivative control to any of the above controller types, we reach the same steady-state equation as the derivative of the error vanishes at steady state.

In conclusion, Theorem 1 is compatible with all possible combinations of the proportional-integral-derivative (PID) controller framework, except for the pure derivative control. Furthermore, all other controller types that satisfy the steady-state equation  $F(\mathbf{u}_{\text{ss}}, \boldsymbol{\theta})$  are also compatible with Theorem 1.

### A.6. Adapting the minimizing control framework for classification with softmax and cross-entropy loss

For classification with neural networks, one uses conventionally a one-hot label  $\mathbf{r}_L^{\text{true}}$  in combination with a cross-entropy loss and a softmax layer as output of the network. In our minimizing control framework, the controller drives the network to perfectly match the output target  $\mathbf{r}_L^{\text{true}}$ . However, a softmax layer only results in a one-hot vector if it has an infinite input.

To overcome this challenge, we use soft targets which have  $a$  on its label entry and  $\frac{1-a}{n_L-1}$  on the other entries, with  $a$  close to 1 (e.g., 0.99).

A nice property of the combination of the cross-entropy loss with the softmax output is that the curvature of the exponentials in the softmax cancel out the curvature of the logarithms in the cross-entropy loss, leading to clean gradients that do not saturate easily. However, if our controller uses as control error  $\mathbf{e}(t) = \mathbf{r}_L - \mathbf{r}_L(t)$ , with  $\mathbf{r}_L$  the soft target and  $\mathbf{r}_L(t)$  the softmax output, it cannot make use of the cross-entropy loss to provide clean gradients. To overcome this challenge, we use a linear output layer,  $\mathbf{r}_L$ , and absorb the softmax into the cross-entropy loss, leading to the combined loss:

$$\mathcal{L}^{\text{combined}} = - \sum_{b=1}^B \mathbf{r}_L^{\text{true},(b)T} \log(\text{Softmax}(\mathbf{r}_L^{(b)})). \quad (40)$$

Next, we slightly adapt the minimizing control framework to include this combined loss in a principled way. Currently, the controller dynamics is given by

$$\mathbf{u}(t) = k\mathbf{e}(t) + \mathbf{u}^{\text{int}}(t), \quad \tau_u \frac{d}{dt} \mathbf{u}^{\text{int}}(t) = \mathbf{e}(t) - \alpha \mathbf{u}^{\text{int}}(t) \quad (41)$$

$$\mathbf{e}(t) = \mathbf{r}_L - \mathbf{r}_L(t). \quad (42)$$

Note that  $\mathbf{e}(t)$  can be interpreted as  $-\frac{\partial \mathcal{L}}{\partial \mathbf{r}_L} \Big|_{\mathbf{r}_L = \mathbf{r}_L(t)}$ , with  $\mathcal{L} = \frac{1}{2} \|\mathbf{r}_L - \mathbf{r}_L(t)\|_2^2$ . Now, if we use another loss  $\mathcal{L}$  (i.e.,  $\mathcal{L}^{\text{combined}}$ ), we can generalize the control error  $\mathbf{e}(t)$  to

$$\mathbf{e}(t) = - \frac{\partial \mathcal{L}}{\partial \mathbf{r}_L} \Big|_{\mathbf{r}_L = \mathbf{r}_L(t)}^T. \quad (43)$$

For the cross entropy loss with the softmax, this boils down to

$$\mathbf{e}(t) = - \frac{\partial \mathcal{L}}{\partial \mathbf{r}_L} \Big|_{\mathbf{r}_L = \mathbf{r}_L(t)}^T = \mathbf{p} - \text{Softmax}(\mathbf{r}_L(t)), \quad (44)$$

with  $\mathbf{p}$  the soft target.

## B. Stability analysis

In this section, we build further upon the stability analysis results of [Meulemans et al. \(2021\)](#), which are applicable to Strong-DFC with minimal adaptations.

As a first step, we investigate a pure integral controller ( $k = 0$ ) and assume a separation of timescales  $\tau_v \ll \tau_u$ , i.e., we can replace the network dynamics by a deterministic function and only retain the controller dynamics. This provides us with the following controller dynamics:

$$\tau_u \dot{\mathbf{u}}(t) = \mathbf{e}(\mathbf{u}(t)) - \alpha \mathbf{u}(t), \quad (45)$$

where the error,  $\mathbf{e}$ , can be seen as a deterministic function of the control input,  $\mathbf{u}(t)$  (for simplicity, we omit the dependence on the network parameters and input, as they are considered fixed during the controller dynamics). To investigate the local stability at steady state, we consider the eigenvalues of the Jacobian matrix of these dynamics evaluated at steady state. Using Eq. (23), we have that:

$$\frac{\partial}{\partial \mathbf{u}} (\mathbf{e}(\mathbf{u}) - \alpha \mathbf{u}) \Big|_{\mathbf{u} = \mathbf{u}_{\text{ss}}} = - (J_{\text{ss}} Q + \alpha I). \quad (46)$$

Hence, the condition for local stability around the steady state is that all eigenvalues of  $-(J_{\text{ss}} Q + \alpha I)$  should be negative, resulting in Condition 2.

When considering the proportional integral controller without a separation of timescales, the local stability condition becomes harder to interpret. Following a similar derivation as proposed by [Meulemans et al. \(2021\)](#), we reach the following Jacobian of the network-controller dynamics:

$$A_{\text{PI}} = - \begin{bmatrix} \frac{1}{\tau_v} (I - W \phi^\theta) & -\frac{1}{\tau_v} Q \\ \frac{1}{\tau_u} S \phi^\theta - \frac{k}{\tau_v} S \phi^\theta (I - W \phi^\theta) & \frac{k}{\tau_v} S \phi^\theta Q \frac{\tilde{\alpha}}{\tau_u} I \end{bmatrix} \quad (47)$$



with  $\tilde{\tau}_u$ ,  $\frac{\tau_u}{1+\alpha k}$ ,  $\tilde{\alpha}$ ,  $\frac{\alpha}{1+\alpha k}$ ,  $\phi^\theta$ ,  $\left. \frac{\partial \phi(\mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{v}=\mathbf{v}_{ss}}$ , and

$$W = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ W_2 & 0 & 0 & \dots & 0 \\ 0 & W_3 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & 0 \\ 0 & 0 & \dots & W_L & 0 \end{bmatrix} \quad (48)$$

$$S = [0 \quad \dots \quad 0 \quad I] \quad (49)$$

The condition for local stability is now that all eigenvalues of  $A_{PI}$  are negative. However, this condition provides no straightforward interpretations for defining conditions on the feedback weights,  $Q$ , such that local stability is achieved. One can apply Gershgoring's circle theorem for finding sufficient conditions on  $J_{ss}$  and  $Q$  to ensure local stability, however, the resulting conditions are too conservative and hence not usable for designing feedback weight learning rules.

## C. Proofs and discussion for Section 4

### C.1. Stochastic dynamics for the single-phase learning setting

In this section, we provide a detailed description of the stochastic dynamics used for training the forward and feedback weights in a single phase. The network and controller obey the following dynamics:

$$\tau_v \frac{d}{dt} \mathbf{v}_i(t) = -\mathbf{v}_i(t) + W_i \phi(\mathbf{v}_{i-1}(t)) + Q_i \mathbf{u}(t) + \sigma \epsilon_i(t) \quad (50)$$

$$\tau_u \frac{d}{dt} \mathbf{u}^{\text{int}}(t) = \mathbf{r}_L - \mathbf{r}_L(t) - \alpha \mathbf{u}^{\text{int}}(t) \quad (51)$$

$$\mathbf{u}(t) = k(\mathbf{r}_L - \mathbf{r}_L(t)) + \mathbf{u}^{\text{int}}(t) \quad (52)$$

$$(53)$$

with  $\sigma^2$  the noise variance and  $\epsilon(t)$  exponentially filtered white noise, with time constant  $\tau_\epsilon$ :

$$\epsilon_i(t) = \frac{1}{\tau_\epsilon} \int_0^t \exp\left(-\frac{1}{\tau_\epsilon}(t-t^\theta)\right) \xi_i(t^\theta) dt^\theta, \quad (54)$$

where  $\xi$  is white noise. The time constant  $\tau_\epsilon$  determines the auto-correlation of  $\epsilon$ . For simplicity, we take that the noise  $\epsilon$  enters solely in the feedback compartment:

$$\mathbf{v}_i^{\text{fb}}(t) = Q_i \mathbf{u}(t) + \sigma \epsilon_i(t). \quad (55)$$

However, all the arguments and theorems of the following sections also apply for noise entering all compartments.

Now, we define the following error-based learning rule for the forward weights and anti-Hebbian learning rule for the feedback weights:

$$\tau_W \frac{d}{dt} W_i(t) = \left( \phi(\mathbf{v}_i(t)) - \phi(W_i \mathbf{r}_{i-1}(t)) \right) \bar{\mathbf{r}}_i^T \quad (56)$$

$$\tau_Q \frac{d}{dt} Q_i(t) = - \left( 1 + \frac{\tau_v}{\tau_\epsilon} \right)^{L-i} \mathbf{v}_i^{\text{fb}}(t) \tilde{\mathbf{u}}(t)^T - \beta Q_i \quad (57)$$

$$\tilde{\mathbf{u}}(t) = \mathbf{u}(t) - \bar{\mathbf{u}}(t) \quad (58)$$

$$\bar{\mathbf{u}}(t) = \frac{1}{\tau_f} \int_0^t \exp\left(-\frac{1}{\tau_f}(t-t^\theta)\right) \mathbf{u}(t^\theta) dt^\theta \quad (59)$$

$$\bar{\mathbf{r}}_i(t) = \frac{1}{\tau_f} \int_0^t \exp\left(-\frac{1}{\tau_f}(t-t^\theta)\right) \mathbf{r}_i(t^\theta) dt^\theta \quad (60)$$

with  $\tau_f$  the filtering time constant used to high-pass filter  $\mathbf{u}$  and low-pass filter  $\mathbf{r}_i$ .

## C.2. Intuition behind the feedback weight learning rule

Here we expand on the intuition behind our feedback weight learning rule. The central technique, inspired by the weight mirroring method (Akrouf et al., 2019) and DFC (Meulemans et al., 2021), is to inject white noise in the network that carries information about the network Jacobian,  $J_{ss}$ , towards the output. If we now measure the correlation of the output noise w.r.t. the injected noise in the layers, we can recover the information on  $J_{ss}$ . Below, we summarize the intuitive explanation on a normal feedforward neural network, provided by Meulemans et al. (2021). We end with explaining how to go from this intuitive explanation to our single-phase feedback weight plasticity, and how it differs from the original feedback weight plasticity of DFC.

Now, let us consider a normal feedforward neural network  $\mathbf{r}_i = \phi(\mathbf{v}_i) = \phi(W_i \mathbf{r}_{i-1})$ . We perturb each layer with white noise  $\sigma \boldsymbol{\xi}$ , which is propagated forward through the network:

$$\tilde{\mathbf{v}}_i = W_i \phi(\tilde{\mathbf{v}}_{i-1}) + \sigma \boldsymbol{\xi}_i \quad (61)$$

$$\tilde{\mathbf{r}}_0 = \mathbf{r}_0 \quad (62)$$

Assuming  $\sigma$  is small, we can approximate the network output as:

$$\tilde{\mathbf{r}}_L \approx \mathbf{r}_L + \sigma J \boldsymbol{\xi}, \quad (63)$$

with  $J = \left[ \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1}, \dots, \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L} \right]$  and  $\boldsymbol{\xi}$  the concatenation of all  $\boldsymbol{\xi}_i$ . Now, we can use an output error  $\mathbf{e} = \mathbf{r}_L - \tilde{\mathbf{r}}_L$  to extract the noise, and use this in a feedback learning rule  $\Delta Q = -\boldsymbol{\xi} \mathbf{e}^T - \beta Q$ . In expectation, this update rule results in

$$E[\Delta Q] = \sigma^2 J^T - \beta Q. \quad (64)$$

Hence, we see that this simple update rule drives the feedback weights to align with  $J^T$ .

Although the above explanation grasps the intuition behind the feedback learning rule (9), there are some important differences when we consider the continuous single-phase plasticity dynamics.

1. The pre-synaptic plasticity signal of the feedback weight update (9) is the high-pass filtered control signal  $\mathbf{u}_{\text{hp}}$  instead of the output noise we used before. When training the feedback and forward weights simultaneously,  $\mathbf{u}$  consists of two parts: (i) the control signal used to drive the network in expectation towards the output target  $\mathbf{r}_L$ , and (ii), the output noise fluctuations integrated over time by the controller. As we are interested in the noise fluctuations, we need to extract them, which we do by using the high-pass filtered version of  $\mathbf{u}$ .
2. The postsynaptic plasticity signal is the activity in the feedback compartment  $\mathbf{v}_i^{\text{fb}} = Q_i \mathbf{u} + \epsilon_i$ , instead of purely the noise. As  $Q_i \mathbf{u}$  will correlate with  $\mathbf{u}_{\text{hp}}$ , and  $\mathbf{u}_{\text{hp}}$  contains the integrated noise fluctuations over time, instead of the direct noise fluctuations, the feedback weights  $Q$  will not align exactly with  $J^T$ , but with  $J^T M$  instead, where  $M$  is a positive definite matrix.
3. To incorporate noise in the continuous dynamics of Strong-DFC, we need to use stochastic differential equations (SDEs), instead of discrete noise perturbations. Unlike discrete feedforward neural networks, the network dynamics do not propagate the layer activity instantaneously towards the output, but have a delay proportional to  $\tau_v$ . As white noise has no correlation in time, naively correlating the output noise fluctuations with the current layer noise would result in a signal of zero due to the time delays. In order to have correlations over time, we use exponentially filtered white noise  $\epsilon$  with time constant  $\tau_\epsilon$ , instead of pure white noise  $\boldsymbol{\xi}$ . As the time delay between the output noise fluctuations and the layer noise increases for upstream layers, we need to compensate for the decreasing correlation signal by scaling the plasticity updates with  $\left(1 + \frac{\tau_v}{\tau_\epsilon}\right)^{L-i}$ , as is done in Eq. (9).

Although our single-phase plasticity dynamics (9) are closely related to the two-phase plasticity dynamics of DFC introduced by Meulemans et al. (2021), and these authors discussed briefly the possibility of having single-phase plasticity dynamics, our plasticity dynamics introduce some important advances, both in theory and practice.

1. Our minimizing control framework explicitly motivates learning the forward and feedback weights simultaneously, as the feedback weights are trained to satisfy Condition 1, which is evaluated at the steady state of the network for training

the forward weights. By contrast, the theory for DFC in the weak-nudging regime requires the feedback weights to be trained at the network steady state without influence from the controller, and therefore different from the network steady state used for training the forward weights (Meulemans et al., 2021).

2. In their theoretical results on the feedback learning rule, Meulemans et al. (2021) require that the controller leakage  $\alpha$  is big. However, this cannot be combined with learning the feedback and feedforward weights simultaneously, as for training the feedforward weights, we need  $\alpha \rightarrow 0$ . In our new theoretical results, we overcome this important challenge.
3. In their theoretical results on the feedback learning rule, Meulemans et al. (2021) assume instant network dynamics ( $\tau_v \rightarrow 0$ ). However, in practice, a finite network time constant is used, giving rise to delays in the propagation of noise fluctuations to the output, leading to weaker correlation signals for upstream layers. Our new theory embraces finite time constants and investigates in detail the required relation between the various time constants  $\tau_v$ ,  $\tau_\epsilon$ ,  $\tau_u$ , and  $\tau_f$ . As a result, we introduce the scaling term  $\left(1 + \frac{\tau_v}{\tau_\epsilon}\right)^{L-i}$  which improves the feedback learning significantly.

### C.3. Proof of Theorem 3

In the following, we prove Theorem 3, which we state in its full form below.

**Theorem S4.** *Assume a separation of timescales  $\tau_v, \tau_\epsilon \ll \tau_u \ll \tau_f \ll \tau_Q$ , integral control ( $k = 0$ ), stable network dynamics and small noise perturbations  $\sigma \ll \|\mathbf{v}_i\|$ . Then, for a fixed data sample, the feedback plasticity dynamics (9) let the first moment of  $Q$  converge approximately towards the following configuration that satisfies Conditions 1 and 2:*

$$\mathbb{E}[Q_{\text{ss}}] \approx J_{\text{ss}}^T M, \quad (65)$$

with  $M$  a positive definite symmetric matrix.

*Proof.* We consider the noisy dynamics defined in Eq. (50-52), with  $k = 0$ . Let us define  $\bar{\mathbf{v}}_{\text{ss}}$  and  $\bar{\mathbf{u}}_{\text{ss}}$  as the steady state of the mean trajectory of the noisy dynamics, and  $\tilde{\mathbf{v}}, \mathbf{v} - \bar{\mathbf{v}}_{\text{ss}}$  and  $\tilde{\mathbf{u}}, \mathbf{u} - \bar{\mathbf{u}}_{\text{ss}}$  as the (noisy) perturbations around the mean trajectories. As  $\sigma \ll \|\mathbf{v}_i\|$  and the dynamics are stable (hence, contracting), we have that  $\|\tilde{\mathbf{v}}\| \ll \|\mathbf{v}\|$  and  $\|\tilde{\mathbf{u}}\| \ll \|\mathbf{u}\|$ , allowing us to do a first-order Taylor expansion of the dynamics, which for simplicity we assume to be exact:

$$\tau_v \frac{d}{dt} \tilde{\mathbf{v}}(t) = -\bar{\mathbf{v}}_{\text{ss}} + W\phi(\bar{\mathbf{v}}_{\text{ss}}) + U\mathbf{r}_0 + Q\bar{\mathbf{u}}_{\text{ss}}(t) - \tilde{\mathbf{v}}(t) + W\phi^\theta \tilde{\mathbf{v}}(t) + Q\tilde{\mathbf{u}}(t) + \sigma\boldsymbol{\epsilon}(t) \quad (66)$$

$$= -\tilde{\mathbf{v}}(t) + W\phi^\theta \tilde{\mathbf{v}}(t) + Q\tilde{\mathbf{u}}(t) + \sigma\boldsymbol{\epsilon}(t) \quad (67)$$

$$\tau_u \frac{d}{dt} \tilde{\mathbf{u}}(t) = \mathbf{r}_L - S\phi(\bar{\mathbf{v}}_{\text{ss}}) - \alpha\bar{\mathbf{u}}_{\text{ss}} - S\phi^\theta \tilde{\mathbf{v}}(t) - \alpha\tilde{\mathbf{u}}(t) \quad (68)$$

$$= -S\phi^\theta \tilde{\mathbf{v}}(t) - \alpha\tilde{\mathbf{u}}(t) \quad (69)$$

where we define  $\phi^\theta, \left. \frac{\partial \phi(\mathbf{v})}{\partial \mathbf{v}} \right|_{\mathbf{v}=\bar{\mathbf{v}}_{\text{ss}}}$ ,

$$W, \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ W_2 & 0 & 0 & \dots & 0 \\ 0 & W_3 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & W_L & 0 \end{bmatrix} \quad (70)$$

$$U, [W_1^T \ 0 \ \dots \ 0]^T \quad (71)$$

$$S, [0 \ \dots \ 0 \ I] \quad (72)$$

and use that  $-\bar{\mathbf{v}}_{\text{ss}} + W\phi(\bar{\mathbf{v}}_{\text{ss}}) + U\mathbf{r}_0 + Q\bar{\mathbf{u}}_{\text{ss}}(t) = 0$  and  $\mathbf{r}_L - S\phi(\bar{\mathbf{v}}_{\text{ss}}) - \alpha\bar{\mathbf{u}}_{\text{ss}} = 0$  at the steady state of the mean trajectory. Similar to  $\Delta \mathbf{v}_i$ , we define  $\Delta \tilde{\mathbf{v}}_i, \tilde{\mathbf{v}}_i - W_i \phi_{i-1}^\theta \tilde{\mathbf{v}}_{i-1}$  with  $\phi_{i-1}^\theta, \left. \frac{\partial \phi(\mathbf{v}_i)}{\partial \mathbf{v}_i} \right|_{\mathbf{v}=\bar{\mathbf{v}}_{\text{ss},i}}$ . As  $\Delta \tilde{\mathbf{v}}_1 = \tilde{\mathbf{v}}_1$  and  $\tilde{\mathbf{v}}_i = \Delta \tilde{\mathbf{v}}_i + W_i \phi_{i-1}^\theta$ , we have that

$$S\phi^\theta \tilde{\mathbf{v}} = \phi_L^\theta \tilde{\mathbf{v}}_L = J_{\text{ss}} \Delta \tilde{\mathbf{v}}, \quad (73)$$

with  $J_{ss}$ ,  $[\frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1}, \dots, \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L}]$  evaluated at  $\bar{\mathbf{v}}_{ss}$ . Using  $\frac{d}{dt} \Delta \tilde{\mathbf{v}} = (I - W\phi^\theta) \frac{d}{dt} \tilde{\mathbf{v}}$  and  $\gamma(t)$ ,  $[\Delta \tilde{\mathbf{v}}(t)^T, \tilde{\mathbf{u}}(t)^T]^T$  to simplify notation we get

$$\frac{d}{dt} \gamma(t) = - \underbrace{\begin{bmatrix} \frac{1}{\tau_v} (I - W\phi^\theta) & -\frac{1}{\tau_v} (I - W\phi^\theta) Q \\ \frac{1}{\tau_u} J_{ss} & \frac{\alpha}{\tau_u} I \end{bmatrix}}_{\triangleq A} \gamma(t) + \underbrace{\begin{bmatrix} \frac{\sigma}{\tau_v} (I - W\phi^\theta) \\ 0 \end{bmatrix}}_{\triangleq B} \epsilon(t) \quad (74)$$

We solve this set of linear time-invariant differential equations using the variation of constants method (Särkkä & Solin, 2019), taking  $\tilde{\mathbf{v}}(t_0) = 0$  and  $\tilde{\mathbf{u}}(t_0) = 0$  (i.e., we consider that the network and controller dynamics are already converged at  $t_0$ ).

$$\gamma(t) = \int_{t_0}^t \exp(-A(t-t^\theta)) B \epsilon(t^\theta) dt^\theta. \quad (75)$$

Now that we solved the differential equations for  $\tilde{\mathbf{v}}$  and  $\tilde{\mathbf{u}}$ , we turn our attention to the feedback weight plasticity rule (9). We start by investigating  $E[\mathbf{v}^{fb}(t) \tilde{\mathbf{u}}(t)^T]$ . As the noise  $\epsilon$  enters through the apical compartment, we have that:

$$\mathbf{v}^{fb}(t) = Q\mathbf{u}(t) + \sigma\epsilon(t) = Q\bar{\mathbf{u}}_{ss} + Q\tilde{\mathbf{u}}(t) + \sigma\epsilon(t) \quad (76)$$

$$E[\mathbf{v}^{fb}(t) \tilde{\mathbf{u}}(t)^T] = Q\bar{\mathbf{u}}_{ss} E[\tilde{\mathbf{u}}(t)^T] + QE[\tilde{\mathbf{u}}(t) \tilde{\mathbf{u}}(t)^T] + \sigma E[\epsilon(t) \tilde{\mathbf{u}}(t)^T] \quad (77)$$

The first term of Eq. (77) is equal to zero, as  $\epsilon$  and consequently  $\tilde{\mathbf{u}}$  are zero-mean. We now turn our attention to the third term. Using  $S_u$ ,  $\begin{bmatrix} 0 & I \end{bmatrix}$  and  $\tilde{\mathbf{u}} = S_u \gamma$ , we have that:

$$E[\tilde{\mathbf{u}}(t) \epsilon(t)^T] = S_u E[\gamma(t) \epsilon(t)^T] = S_u \int_{t_0}^t \exp(-A(t-t^\theta)) B E[\epsilon(t^\theta) \epsilon(t)^T] dt^\theta \quad (78)$$

$$= \frac{1}{2\tau_\epsilon} S_u \int_{t_0}^t \exp(-(A + \frac{1}{\tau_\epsilon} I)(t-t^\theta)) B dt^\theta, \quad (79)$$

for which we used  $E[\epsilon(t^\theta) \epsilon(t)^T] = \frac{1}{2\tau_\epsilon} \exp(\frac{1}{\tau_\epsilon} |t-t^\theta|)$ . Using  $\int_0^T \exp(At) dt = (\exp(AT) - I) A^{-1}$  and assuming stability and  $t-t_0 \gg \tau_u$ , which is justified by the condition  $\tau_Q \gg \tau_u$ , we can solve the previous integral:

$$E[\tilde{\mathbf{u}}(t) \epsilon(t)^T] = \frac{1}{2\tau_\epsilon} S_u (A + \frac{1}{\tau_\epsilon} I)^{-1} B. \quad (80)$$

Using the solution for the inverse of  $2 \times 2$  block matrices of Lu & Shiou (2002), we get

$$E[\tilde{\mathbf{u}}(t) \epsilon(t)^T] = \quad (81)$$

$$-\frac{1}{2\tau_\epsilon} \left( \frac{1}{\tau_\epsilon} + \frac{\alpha}{\tau_u} \right)^{-1} \frac{1}{\tau_u} J_{ss} \left( \frac{1}{\tau_\epsilon} I + \frac{1}{\tau_v} (I - W\phi^\theta) + \left( \frac{1}{\tau_\epsilon} + \frac{\alpha}{\tau_u} \right)^{-1} \frac{1}{\tau_v} (I - W\phi^\theta) Q \frac{1}{\tau_u} J_{ss} \right)^{-1} \frac{\sigma}{\tau_v} (I - W\phi^\theta) \quad (82)$$

Now, using the condition  $\tau_v, \tau_\epsilon \ll \tau_u$ , we can approximate the above with:

$$E[\tilde{\mathbf{u}}(t) \epsilon(t)^T] \approx -\frac{\sigma}{2\tau_u} J_{ss} \left( \left( 1 + \frac{\tau_v}{\tau_\epsilon} \right) I - W\phi^\theta \right)^{-1} (I - W\phi^\theta) \quad (83)$$

Using Lemma S5, we have that:

$$E[\tilde{\mathbf{u}}(t) \epsilon(t)^T] \approx -\frac{\sigma}{2\tau_u} \left[ \left( 1 + \frac{\tau_v}{\tau_\epsilon} \right)^L \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1} \quad \left( 1 + \frac{\tau_v}{\tau_\epsilon} \right)^{(L-1)} \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_2} \quad \dots \quad \left( 1 + \frac{\tau_v}{\tau_\epsilon} \right)^1 \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L} \right] \Bigg|_{\mathbf{v}=\bar{\mathbf{v}}_{ss}} \quad (84)$$

Now we turn our attention towards the second term of Eq. (77). As  $E[\tilde{\mathbf{u}}(t) \tilde{\mathbf{u}}(t)^T]$  is a covariance matrix, it is positive semi-definite.



Combining Eq. (9) with Eq. (81) and the above insights, we get

$$\tau_Q \frac{d}{dt} \mathbb{E}[Q(t)], \quad \tau_Q \frac{d}{dt} \bar{Q}(t) = \frac{\sigma^2}{2\tau_u} J_{ss}^T - \bar{Q}(\tilde{M} + \beta I), \quad (85)$$

with  $\tilde{M}$  a positive semi-definite matrix. We used the fact that the layer-wise scaling with  $\left(1 + \frac{\tau_v}{\tau_\epsilon}\right)^{L-i}$  in Eq. (9) cancels out the scaling factors  $\left(1 + \frac{\tau_v}{\tau_\epsilon}\right)^{(L-i)}$  in Eq. (81). Using  $M$ ,  $\tilde{M} + \beta I$ , the stable point of the above dynamics is

$$\bar{Q}_{ss} = \frac{\sigma^2}{\tau_u} J_{ss}^T M^{-1}. \quad (86)$$

As  $M$  is positive definite, the dynamics are stable and the above steady state is reached. Furthermore, as  $J_{ss}^T M^{-1}$  is a permutation of the columns of  $J_{ss}^T$ , it satisfies Condition 1; and as  $M^{-1}$  is positive definite, it satisfies Condition 2.

Now the only thing left to show is that  $\mathbf{u}_{hp} = \mathbf{u} - \mathbf{u}_{lp} \stackrel{?}{=} \tilde{\mathbf{u}}$ , with  $\mathbf{u}_{lp}$  the low-pass filtered version of  $\mathbf{u}$ :

$$\mathbf{u}_{lp}(t) = \frac{1}{\tau_f} \int_{t_0}^t \exp\left(-\frac{1}{\tau_f}(t-t^\theta)\right) \mathbf{u}(t^\theta) dt^\theta. \quad (87)$$

As we have that  $\tau_f \gg \tau_u, \tau_\epsilon, \tau_v$ , the low-pass filtering will remove all fluctuations and we have that  $\mathbf{u}_{lp} = \bar{\mathbf{u}}_{ss}$ . Consequently, we have that  $\mathbf{u}_{hp} = \mathbf{u} - \bar{\mathbf{u}}_{ss} = \tilde{\mathbf{u}}$ , thereby concluding the proof.  $\square$

**Lemma S5.** For some scalar  $a \neq 0$ , we have that

$$J(aI - W\phi^\theta)^{-1}(I - W\phi^\theta) = \left[ a \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1} \quad a^{(L-1)} \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_2} \quad \dots \quad a \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L} \right] \Big|_{\mathbf{v}=\mathbf{v}^\theta} \quad (88)$$

with  $J$  and  $\phi^\theta$  evaluated at some activity pattern  $\mathbf{v}^\theta$ .

*Proof.* First, we note that  $(aI - W\phi^\theta)$  is of full rank for  $a \neq 0$ , as it is a triangular matrix with  $a$  on its diagonal. The inverse of  $(aI - W\phi^\theta)$  is given by:

$$(aI - W\phi^\theta)^{-1} = \begin{bmatrix} a^{-1}I & 0 & 0 & \dots & 0 \\ a^{-2}J_{2,1} & a^{-1}I & 0 & \dots & 0 \\ a^{-3}J_{3,1} & a^{-2}J_{3,2} & a^{-1}I & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a^{-L}J_{L,1} & \dots & \dots & a^{-2}J_{L,L-1} & a^{-1}I \end{bmatrix} \quad (89)$$

with  $J_{i,j} = \frac{\partial \mathbf{v}_i}{\partial \mathbf{v}_j}$ . One can easily verify this by computing  $(aI - W\phi^\theta)(aI - W\phi^\theta)^{-1} = I$  upon noting that  $J_{i,i-1} = W_i \phi_{i-1}^\theta$ . Next, we have that:

$$(aI - W\phi^\theta)^{-1}(I - W\phi^\theta) = \begin{bmatrix} a^{-1}I & 0 & 0 & \dots & 0 \\ (a^{-2} - a^{-1})J_{2,1} & a^{-1}I & 0 & \dots & 0 \\ (a^{-3} - a^{-2})J_{3,1} & (a^{-2} - a^{-1})J_{3,2} & a^{-1}I & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ (a^{-L} - a^{-(L-1)})J_{L,1} & \dots & \dots & (a^{-2} - a^{-1})J_{L,L-1} & a^{-1}I \end{bmatrix} \quad (90)$$

Using  $J = \left[ \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1}, \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_2}, \dots, \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L} \right]$ , we have

$$J(aI - W\phi^\theta)^{-1}(I - W\phi^\theta) = \left[ a \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_1} \quad a^{(L-1)} \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_2} \quad \dots \quad a \frac{\partial \mathbf{r}_L}{\partial \mathbf{v}_L} \right], \quad (91)$$

thereby concluding the proof.  $\square$

#### C.4. Noise can bias the feedforward weight updates

As observed by Meulemans et al. (2021), when we inject noise into the dynamics while training the forward weights, the resulting noise correlations can bias the forward weight updates. This is a general issue for methods using error feedback for learning while having realistic noise dynamics, as discussed in more detail in App. C.4 in Meulemans et al. (2021). Here, we investigate how these issues manifest themselves in Strong-DFC and how we can solve them.

The forward weight update is given by Eq. (4). If the noise fluctuations in the postsynaptic term  $\phi(\mathbf{v}_i(t)) - \phi(W_i \mathbf{r}_{i-1}(t))$  are correlated to the noise fluctuations in the presynaptic term  $\mathbf{r}_{i-1}(t)$ , the expected forward weight update contains this extra correlation term and is hence biased.

The difference  $\phi(\mathbf{v}_i(t)) - \phi(W_i \mathbf{r}_{i-1}(t))$  is mainly determined by the feedback input  $Q\mathbf{u}(t)$ . Using the notation of Section C.3, we divide the control signal of the converged dynamics in two parts:  $\mathbf{u}(t) = \mathbf{u}_{\text{ss}} + \tilde{\mathbf{u}}(t)$ , with  $\mathbf{u}_{\text{ss}}$  the part of the control signal useful for learning the forward weights, and  $\tilde{\mathbf{u}}$  the noise fluctuations that are used for learning the feedback weights. These noise fluctuations are a function of the noise fluctuations of all network layers (e.g., see Eq. 75). Hence, the noise fluctuations of  $\phi(\mathbf{v}_i(t)) - \phi(W_i \mathbf{r}_{i-1}(t))$  will contain a part of the noise fluctuations of  $\mathbf{r}_{i-1}(t)$ , therefore causing the pre- and post-synaptic term in the plasticity rule to be correlated.

A simple approach to remove the noise correlation in the plasticity rule for the forward weights, is either to remove the noise fluctuations in the pre- or post-synaptic term, or both. We choose to remove the noise fluctuations in the pre-synaptic term, by using a low-pass filtered version of  $\mathbf{r}_{i-1}$ :

$$\bar{\mathbf{r}}_i(t) = \frac{1}{\tau_f} \int_0^t \exp\left(-\frac{1}{\tau_f}(t-t^\theta)\right) \mathbf{r}_i(t^\theta) dt^\theta, \quad (92)$$

with  $\tau_f$  the filtering time constant. This low-pass filtering can be interpreted as a leaky biophysical mechanism that accumulates the pre-synaptic input used for the plasticity of the synapse. As the pre-synaptic term of the forward plasticity dynamics no longer contains the noise fluctuations, the updates will no longer be biased.

## D. Simulation and algorithms

### D.1. Simulating the dynamics

We use the Euler-Maruyama method to simulate the stochastic differential equations (SDE), which simulates a (non)linear SDE

$$\frac{d}{dt} \mathbf{x}(t) = \mathbf{f}(\mathbf{x}, t) + \mathbf{g}(\mathbf{x}, t) \boldsymbol{\xi}(t) \quad (93)$$

as follows:

$$\mathbf{x}[m+1] = \mathbf{x}[m] + \mathbf{f}(\mathbf{x}[m], m\Delta t) \Delta t + \mathbf{g}(\mathbf{x}[m], m\Delta t) \sqrt{\Delta t} \Delta \boldsymbol{\beta}_m, \quad (94)$$

with  $\Delta \boldsymbol{\beta}$  a Gaussian random variable with distribution  $\mathcal{N}(\mathbf{0}, I)$ . Note that the noise term has stepsize  $\sqrt{\Delta t}$  instead of  $\Delta t$ , as Brownian motion has a variance of  $t$  and not  $t^2$ .

Applying the Euler-Maruyama method to the SDEs of Section C.1 defining the neural and controller dynamics gives us the following discrete update equations:

$$\boldsymbol{\epsilon}_i[m+1] = \boldsymbol{\epsilon}_i[m] + \frac{1}{\tau_e} \left( -\boldsymbol{\epsilon}_i[m] \Delta t + \sqrt{\Delta t} \Delta \boldsymbol{\beta}_{i,m} \right) \quad (95)$$

$$\mathbf{v}_i^{\text{fb}}[m] = Q_i \mathbf{u}[m] + \sigma \boldsymbol{\epsilon}_i[m] \quad (96)$$

$$\mathbf{u}^{\text{int}}[m+1] = \mathbf{u}^{\text{int}}[m] + (\mathbf{e}[m] - \alpha \mathbf{u}^{\text{int}}[m]) \frac{\Delta t}{\tau_u} \quad (97)$$

$$\mathbf{u}[m] = k \mathbf{e}[m] + \mathbf{u}^{\text{int}}[m] \quad (98)$$

$$\mathbf{v}_i[m+1] = \mathbf{v}_i[m] + (-\mathbf{v}_i[m] + W_i \phi(\mathbf{v}_{i-1}[m]) + \mathbf{v}_i^{\text{fb}}[m]) \frac{\Delta t}{\tau_v} \quad (99)$$

And for the weight updates:

$$\bar{\mathbf{u}}[m+1] = \bar{\mathbf{u}}[m] + \frac{1}{\tau_f}(-\bar{\mathbf{u}}[m] + \mathbf{u}[m])\Delta t \quad (100)$$

$$\bar{\mathbf{r}}_i[m+1] = \bar{\mathbf{r}}_i[m] + \frac{1}{\tau_f}(-\bar{\mathbf{r}}_i[m] + \mathbf{r}_i[m])\Delta t \quad (101)$$

$$\Delta W_i = \frac{\Delta t}{\tau_W} \sum_m (\mathbf{r}_i[m] - \phi(W_i \mathbf{r}_{i-1}[m])) \bar{\mathbf{r}}_{i-1}[m]^T \quad (102)$$

$$\Delta Q_i = \frac{\Delta t}{\tau_Q} \left(1 + \frac{\tau_v}{\tau_\epsilon}\right)^{L-i} \sum_m -\mathbf{v}_i^{\text{fb}}[m](\mathbf{u}[m] - \bar{\mathbf{u}}[m])^T - M_{\max} \frac{\Delta t}{\tau_Q} \beta Q_i \quad (103)$$

with  $M_{\max}$  the total number of simulation steps.

## D.2. Pseudocode for the Strong-DFC algorithm

Algorithm 1 provides the pseudocode for the Strong-DFC algorithm, making use of the discretized dynamics provided in the previous section. We use the Euler-Maruyama discretization of the SDEs with some slight modifications.

First, in order to have direct control over the hyperparameter  $\tilde{\alpha}$ , we rewrite the controller dynamics as follows:

$$\mathbf{u}(t) = \mathbf{u}^{\text{int}}(t) + k\mathbf{e}(t), \quad \tau_u \frac{d}{dt} \mathbf{u}^{\text{int}}(t) = \mathbf{e}(t) - \tilde{\alpha} \mathbf{u}(t). \quad (104)$$

Note that the underlying dynamics for  $\mathbf{u}(t)$  are exactly the same as in Eq. (2).

In order to better incorporate the layered structure of the network into the discrete network dynamics, we update a layer at timestep  $m+1$  with the previous layer at timestep  $m+1$  instead of  $m$ :  $\mathbf{v}_i^{\text{ff}}[m+1] = W_i \phi(\mathbf{v}_{i-1}[m+1]) + \mathbf{b}_i$ , with  $\mathbf{b}_i$  the bias parameters of layer  $i$ . For small stepsizes  $\Delta t$  this has almost no effect, but for bigger stepsizes it has the benefit of better reflecting the layered structure.

In order to better incorporate the controller loop into the discrete dynamics, we use the updated control signal  $\mathbf{u}[m+1]$  for updating the feedback compartment:  $\mathbf{v}_i^{\text{fb}}[m+1] = Q_i \mathbf{u}[m+1]$ , such that the control error  $\mathbf{e}[m]$  of the previous timestep is used for control instead of the error of two timesteps ago. For small stepsizes  $\Delta t$  this has almost no effect, but for bigger stepsizes it has the benefit of better reflecting the control interaction.

Following the exponential filtering approach for the low-pass filtering of  $\mathbf{r}$  and  $\mathbf{u}$ , we use the current value  $[m+1]$  of  $\mathbf{r}$  and  $\mathbf{u}$  instead of the previous one prescribed by the forward Euler method. This reduces the time delay introduced by the low pass filtering.

## E. Experiments

### E.1. Alignment measures

Below we describe how the measures of Fig. 3 were computed.

**Alignment of Strong-DFC updates with  $\mathcal{H}$**  Fig. 3B quantifies alignment between actual Strong-DFC updates and the gradient of  $\mathcal{H}$ . For this, we vectorize the update matrices and compute the angle between them in degrees.

**Relative strength of feedback control** Fig. 3C quantifies the magnitude of the feedback input compared to the forward drive, and is used to illustrate that our framework allows feedback to have comparable magnitudes to the forward input. It is calculated as follows:

$$\text{ratio}_{\text{fb/ff}} = \frac{\|Q\mathbf{u}\|_F}{\|W\mathbf{r}\|_F}, \quad (105)$$

where  $F$  indicates the Frobenius norm.

**Condition 1** Fig. 3D quantifies the degree to which Condition 1 is satisfied. It is computed through the Frobenius norm of the projection of  $Q$  with the norm of  $Q$ :

$$\text{ratioCon1} = \frac{\|P_{J_{ss}^T} Q\|_F}{\|Q\|_F}. \quad (106)$$

where the projection  $P_{J_{ss}^T} Q$  is given by:

$$P_{J_{ss}^T} Q = J_{ss}^T (J_{ss} J_{ss}^T)^{-1} J_{ss} Q. \quad (107)$$

For an intuition of this metric, please refer to Appendix F.1 in Meulemans et al. (2021).

## E.2. Noise robustness measures

Below we further illustrate the robustness of Strong-DFC weight updates under noisy conditions. Recall that this is a direct consequence of the main novelty of Strong-DFC compared to standard DFC: the use of the true output target  $\mathbf{r}_L = \mathbf{r}_L^{\text{true}}$  throughout training. For this series of experiments (Figure S1) we vary the noise magnitude  $\sigma$  added to the neural dynamics during training of the forward weights for DFC and Strong-DFC. For a fair comparison between both methods, we low-pass filter the presynaptic plasticity signal for debiasing the updates (11). In a student-teacher regression setting with a nonlinear network of size 20-40-5, we found that the performance of Strong-DFC is more robust to a wide range of noise magnitudes compared to standard DFC. For non-noisy dynamics, we observed that standard DFC can descend the loss landscape faster compared to Strong-DFC in this setting, leading to a lower loss, as the Strong-DFC training has not yet fully converged after 500 epochs. For the results of Table 2 in the main manuscript, we used  $\sigma = 0.005$  for both methods.

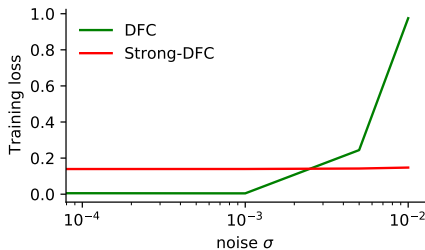


Figure S1. **Noise robustness.** Training loss after 500 epochs (training not fully converged for Strong-DFC) for different noise magnitudes. The DFC results for high noise values are missing since its training was unstable.

## F. Additional Results

### F.1. Training Strong-DFC without re-visiting samples

To investigate whether Strong-DFC requires consistency or re-visiting training samples multiple times to perform well, we performed the following experiment using the student-teacher setting. Using a teacher of size 30-10-10-10-5 and a student of size 30-50-50-50-5, we trained networks for 300 epochs, each containing 500 training samples. In the standard setting, the 500 training samples were identical across epochs. In the modified setting, each training sample was independently drawn from the training distribution, so that no sample was seen twice. This was achieved by feeding the teacher with a new set of randomly-generated input samples at each epoch. Therefore, training consisted of a total of 150000 samples, but in the modified setting all these samples were unique. The results obtained are presented in Table S1.

Table S1. Effect of sample reuse when training Strong-DFC.

Training setting	Train loss	Test loss
Standard	1.945 0.001	2.224 0.002
Without sample reuse	1.964 0.004	2.226 0.006

The results obtained are extremely similar, indicating that Strong-DFC does not necessarily rely on re-visiting training samples many times to perform well. However, it is important to note that Strong-DFC needs i.i.d. samples and will not be

## Minimizing Control for Credit Assignment with Strong Feedback

---

immune to forgetting and interference issues when exposed to correlated sequences of training samples; solving these issues will likely require continual learning techniques, which is out of the scope for this work.



**Algorithm 1** Pseudocode for the Strong-DFC algorithm on a single input sample

```

# Initialize layer activations to the feedforward activations and parameter update buffers to zero
for  $i$  in range(1,  $L$ ) do
     $\mathbf{v}_i[1] = \mathbf{v}_i$ 
     $\mathbf{r}_i[1] = \mathbf{r}_i$ 
     $\Delta W_i = 0$ 
     $\Delta \mathbf{b}_i = 0$ 
     $\Delta Q_i = 0$ 
end for
 $\mathbf{u}^{\text{int}}[1] = 0$ 
for  $m$  in range(1,  $M_{\text{max}}$ ) do
    # Update controller:
    if softmax classification with cross-entropy loss then
         $\mathbf{e}[m] = \mathbf{p} - \text{softmax}(\mathbf{r}_L[m])$  {with  $\mathbf{p}$  the soft target and  $\mathbf{r}_L$  a linear output layer}
    else if regression with L2 loss then
         $\mathbf{e}[m] = \mathbf{r}_L - \mathbf{r}_L[m]$ 
    end if
     $\mathbf{u}^{\text{int}}[m+1] = \mathbf{u}^{\text{int}}[m] + \frac{\Delta t}{\tau_u} (\mathbf{e}[m] - \tilde{\alpha} \mathbf{u}[m])$ 
     $\mathbf{u}[m+1] = \mathbf{u}^{\text{int}}[m+1] + k \mathbf{e}[m]$ 
    # Low-pass filter  $\mathbf{u}$ 
    if  $m == 1$  then
         $\bar{\mathbf{u}}[m+1] = \mathbf{u}[m+1]$  {initialize  $\bar{\mathbf{u}}$ }
    else
         $\bar{\mathbf{u}}[m+1] = \bar{\mathbf{u}}[m] + \frac{1}{\tau_f} (-\bar{\mathbf{u}}[m] + \mathbf{u}[m+1]) \Delta t$ 
    end if
    # Update network:
    for  $i$  in range(1,  $L$ ) do
        # Sample noise to update the exponentially filtered noise variable
         $\Delta \beta_{i,m} \sim \mathcal{N}(0, I)$ 
         $\epsilon_i[m+1] = \epsilon_i[m] + \frac{1}{\tau_e} (-\epsilon_i[m] \Delta t + \sqrt{\Delta t} \Delta \beta_{i,m})$ 
        # Inject noise in the feedback compartment and update the activations
         $\mathbf{v}_i^{\text{fb}}[m+1] = Q_i \mathbf{u}[m+1] + \sigma \epsilon_i[m+1]$ 
         $\mathbf{v}_i^{\text{ff}}[m+1] = W_i \phi(\mathbf{v}_{i-1}[m+1]) + \mathbf{b}_i$ 
         $\mathbf{v}_i[m+1] = \mathbf{v}_i[m] + \frac{\Delta t}{\tau_v} (-\mathbf{v}_i[m] + \mathbf{v}_i^{\text{ff}}[m+1] + \mathbf{v}_i^{\text{fb}}[m+1])$ 
         $\mathbf{r}_i[m+1] = \phi(\mathbf{v}_i[m+1])$ 
        # low-pass filter  $\mathbf{r}_i$ 
         $\bar{\mathbf{r}}_i[m+1] = \bar{\mathbf{r}}_i[m] + \frac{1}{\tau_f} (-\bar{\mathbf{r}}_i[m] + \mathbf{r}_i[m+1]) \Delta t$ 
        # Buffer parameter updates:
         $\Delta W_i = \Delta W_i + (\mathbf{r}_i[m+1] - \phi(W_i \bar{\mathbf{r}}_{i-1}[m+1])) \bar{\mathbf{r}}_{i-1}[m+1]^T$ 
         $\Delta \mathbf{b}_i = \Delta \mathbf{b}_i + \mathbf{r}_i[m+1] - \phi(\mathbf{v}_i^{\text{ff}}[m+1])$ 
         $\Delta Q_i = \Delta Q_i - \left(1 + \frac{\tau_u}{\tau_e}\right)^L \mathbf{v}_i^{\text{fb}}[m] (\mathbf{u}[m+1] - \bar{\mathbf{u}}[m+1])^T - \beta Q_i$ 
    end for
end for
Update parameters with  $\Delta W_i / M_{\text{max}}$ ,  $\Delta \mathbf{b}_i / M_{\text{max}}$  and  $\Delta Q_i / M_{\text{max}}$  using an optimizer of choice
    
```