

Imperial College London
Bioengineering Department

**From Bottom-Up to Top-Down:
A Computational Exploration of Plasticity in
Neural Networks of the Brain**

Klara Kaleb

Thesis submitted for the degree of
Doctor of Philosophy

Statement of originality

I hereby declare that to the best of my knowledge, all work presented in this thesis was performed by me or was referenced otherwise.

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Neural plasticity, the brain’s remarkable capacity for adaptation and learning, is essential for acquiring new skills and knowledge. Computational modelling plays a crucial role in understanding the intricate mechanisms of neural plasticity, bridging the gap between experimental findings and theoretical insights. This thesis explores three computational approaches to address challenges in understanding biological learning:

Network Homeostasis. The first part of this thesis proposes a novel form of inhibitory plasticity that enables global control of network activity, going beyond neuron-centered homeostasis. The model proposed can explain how neural representations can remap while maintaining network stability, both in the hippocampus and the cortex.

Temporal Credit Assignment. The second part of this thesis investigates the mechanisms of a biologically plausible model for rapid motor adaptation. This work demonstrates that feedback control allows accurate and efficient learning in recurrent neural networks, even without a computationally expensive separate relaxation phase.

Spatial Credit Assignment. Using a meta-learning framework, the third part of this thesis explores the space of effective, biologically plausible feedback learning rules in cortical multi-layer networks. A teacher-student network setup is introduced to evaluate credit assignment, leading to the discovery of a novel feedback learning rule that demonstrates successful, biologically plausible learning.

By tackling these challenges, this thesis contributes to a deeper understanding of learning in the brain, paving the way for more effective computational models.

Acknowledgements

Thank you to my supervisor, **Claudia**, for always reminding me of the pleasure of finding things out.

Thank you to my husband, **Tudor**, for always providing a safe space for me to thrive.

Thank you to my dog, **Coco**, for always reminding me to take a break.

Thank you to the **Wellcome Trust**, for their generous funding that made this PhD possible.

Last but not least, thank you to **all my collaborators, family & friends**, for making both research and life fun.

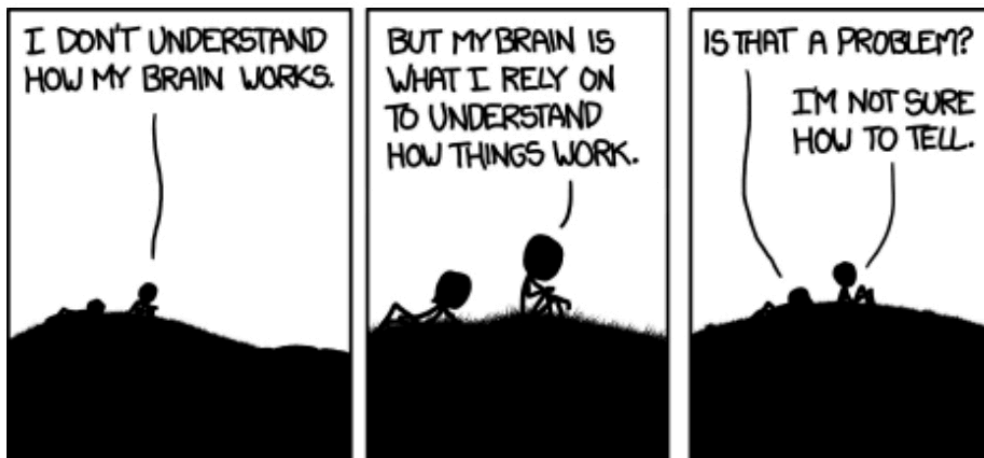


Figure 1: A cartoon humorously addressing the irony of the brain trying to comprehend its own processes, which originally sparked the author's interest in this fascinating topic. source: "Debugger" comic from XKCD.com [link].

Contents

Statement	3
Abstract	3
Acknowledgements	5
1 Introduction	25
1.1 Computational models of biological learning	25
1.2 The trouble with learning in the brain	28
1.3 Thesis structure	30
2 Network-centered homeostasis through inhibition	32
2.1 Abstract	34
2.2 Introduction	35
2.3 Results	38
2.3.1 Input-dependent inhibitory plasticity (IDIP) rule as a homeostatic mechanism	38
2.3.2 The IDIP rule allows for the emergence of active and silent place cells in a hippocampal network	39

2.3.3	The IDIP rule enables rapid homeostatic remapping during active place cells silencing	42
2.3.4	The IDIP rule establishes global E/I balance in recurrent networks . . .	44
2.3.5	The IDIP rule enables maintenance of neural representations	45
2.3.6	The IDIP rule enables memory trace persistence and recall	47
2.4	Discussion	49
2.5	Methods	54
2.5.1	Neuron Model	54
2.5.2	Input-dependent inhibitory plasticity (IDIP)	54
2.5.3	Hippocampal Network Model	55
2.5.4	Recurrent Network Model	57
2.6	Simulations parameters summary	60
2.7	Supplementary figures	63
2.8	Supplementary methods	79
2.8.1	Various learning rule implementations	79
2.8.2	Supplementary parameter tables	81
3	Feedback guides learning in biological circuits	83
3.1	Introduction	85
3.2	Methods	87
3.2.1	Training stages	87
3.2.2	Network architecture	88
3.2.3	Pretraining procedure	89

3.2.4	Adaptation and feedback-driven plasticity rule	89
3.3	Results	90
3.3.1	Recurrent neural networks with feedback control adapt well to task perturbation using feedback control and local learning rules	90
3.3.2	Feedback control approximates the true gradient w.r.t. networks activations during task perturbation	91
3.3.3	Feedback control <i>and</i> local, feedback-driven learning enable rapid network adaptation during persistent task perturbation	91
3.3.4	Feedback control enables accurate recurrent weight adaptation during task perturbation	92
3.3.5	Feedback control enables efficient recurrent weight adaptation during task perturbation	95
3.4	Discussion	96
3.5	Impact statement	98
3.6	Appendix	98
3.6.1	Detailed Methods: Task	98
3.6.2	Detailed Methods: Network	99
3.6.3	Detailed Methods: Pretraining	99
3.6.4	Detailed Methods: Fine-tuning	99
3.6.5	Detailed Methods: Computational requirements	100
3.6.6	Key differences from Feulner et al. (2022)	100
3.6.7	Detailed Methods: Analysis details	100
4	Meta learning feedback plasticity rules	102
4.1	Introduction	104

4.2	Background	106
4.2.1	Learning in Multi-layer Feedforward networks	106
4.2.2	Meta-Learning of <i>Biologically Plausible</i> Learning Rules	108
4.3	Methods	109
4.3.1	The learning task	109
4.3.2	The learning rule formulation	110
4.3.3	Meta-learning setup	112
4.4	Results	112
4.4.1	Underparametrized student-teacher setup reveals a performance gap between backpropagation and feedback alignment	112
4.4.2	Meta-learning identifies a limited set of learning rule parameters for the direct feedback weights	114
4.4.3	Meta-learned learning rule closes the performance gap between backpropagation and direct feedback alignment and aligns the pseudo-gradients with a limited number of examples	115
4.5	Discussion	116
4.6	Appendix	117
4.6.1	Teacher Network Architecture and Task Complexity	117
5	Conclusion	120
5.1	Summary	120
5.2	Future Work	121
5.3	Conclusion	122
	Bibliography	122

List of Tables

2.1	Neuron parameters.	60
2.2	Synapse parameters.	60
2.3	Plasticity parameters.	60
2.4	Simulation parameters.	61
2.5	Hippocampal neural network parameters.	61
2.6	Recurrent neural network parameters.	62
2.7	The parameters used in Figure 2.23B	81
2.8	The parameters used in Figure 2.23A.	82
2.9	The parameters used in Figure 2.24.	82

List of Figures

1	A cartoon humorously addressing the irony of the brain trying to comprehend its own processes, which originally sparked the author’s interest in this fascinating topic. source: “Debugger” comic from XKCD.com [link].	6
2.1	The IDIP rule enables emergence of active and silent place cells in a hippocampal network. A. Inhibitory plasticity rule diagram. An inhibitory neuron (blue) adjusts its inhibitory synaptic weights based on the synaptic input it receives and its target input θ. B. Hippocampal network diagram. The CA3 excitatory neurons (purple) receive place-dependent external current. CA3 neurons project to uniquely tuned CA1 excitatory neurons (green), which are interconnected with CA1 inhibitory neurons (blue). The CA3 to CA1 excitatory synapses are plastic under Hebbian plasticity and the CA1 inhibitory synapses are plastic under IDIP. C-E State of the network during the first lap, before any learning. C. A diagram of a sample microcircuit found in the network in B, consisting of 2 CA1 place cells (1 and 2) with the same spatial tuning, but with cell 2 having stronger place tuning. D. TThe excitatory (purple), inhibitory (blue) and net (green) currents received by cells 1 and 2 during the first lap on the track. The higher peak net current (black dashed line) received by cell 2 reflects its stronger tuning. E. The firing rates of all the CA1 place cells (y-axis) during the first lap (x-axis). All the place cells are active but with varying amplitudes, reflecting differences in their place tuning strengths. F-H The same as in C-E but after learning. Due to increased inhibition, silent (grey triangles) place cells emerge, as the difference in the peak net currents from the first lap is amplified with learning.	41

2.2 Place cells silencing leads to a rapid emergence of an alternative place

map. **A.** Silencing protocol similar to Trouche et al. (2016). All the active cells (green) are tagged, silenced for one lap (yellow background) and released in the subsequent lap. **B-E** Silencing in the network following IDIP. **B.** Network diagram after the exploration phase (100 laps). **C.** Mean firing rate of the network just before, during (yellow background), and after silencing. During silencing there is rapid activation of an alternative place map, consisting of the previously silent cells (grey). After silencing, the original map reactivates (green). **D.** Alternative place map activation just after silencing. **E.** The mean firing rate of the place map before ('OFF') and during ('ON') silencing. Gray circles indicate individual trials. Black circles indicate average over all 20 trials. **F-H** Same as in **B-E** but with inhibitory plasticity turned off during the silencing. **F.** Network diagram after inhibitory and excitatory learning (lap 100) with IDIP turned off during silencing (2). **G.** Without IDIP during the silencing (2) there is low alternative map activation. **H.** Without IDIP during the silencing the mean firing rate of the place map decreases. Gray circles indicate individual trials. Black circles indicate average over all 20 trials. **I** Consolidation protocol similar to Trouche et al. (2016). The established place map is silenced for 4 consecutive laps, and then released for a single testing lap. This is repeated for 10 trials. **J.** Mean firing rate of the active (green) and silent (grey) place cells during the testing laps at each trial. Trial 0 corresponds to the activity of the network just after the exploration phase. 43

- 2.3 IDIP establishes global E/I balance in recurrent networks.**
- A.** A recurrent network of 80 excitatory (white triangles) and 20 inhibitory (blue circles) leaky integrate-and-fire neurons receiving external input from a pool of 100 Poisson excitatory neurons. Only the inhibitory synaptic weights are made plastic with IDIP. **B.** The evolution of the mean excitatory firing rate (black) and the mean inhibitory weight (blue) of the network. The shaded regions correspond to the standard deviation of individual units. IDIP is turned on at 15s. Activity of the network decreases as inhibitory weights increase. **C.** The excitatory firing rate stabilises to a mean value of 5.2 Hz (left). The network displays an asynchronous firing pattern (right). **D.** The range of mean network firing rates IDIP can support, with varying input target value (y-axis) and inhibitory connectivity (x-axis). 'X' marks the parameter combination used in our simulations. **E.** Spike raster plots at the three time points ('X') marked in **B**. The network progresses from high synchronous to low asynchronous firing following the IDIP rule. **F.** Mean membrane currents as a function of excitatory neuron index at three the time points ('X') marked in **B**. 46

- 2.4 IDIP rule enables maintenance of neural representations and memory trace persistence.** **A.** The same network as in Figure 2.3, but with a subset of excitatory neurons (orange) receiving increased external input (grey). **B.** The mean firing rates of the network units, color coded as in A. Bold lines denote the mean and the shaded area indicate STD over 20 trials. **C.** The mean firing rate of the networks without (left) and with (right) increased input to a subset of neurons. Gray circles indicate individual trials and the black circles indicate average over all 20 trials. **D.** Network representation task diagram. We calculate the preservation of the correlation with the initial firing rate rank during the entire course of the simulation. **E.** The Spearman rank correlation coefficient for the network following IDIP (black) and iSTDP (grey). Error bars correspond to STD over 20 trials. **F-H.** Performance of IDIP in an associative memory task. **F.** Associative memory task protocol. Following network stabilization, we increase recurrent excitatory connections between a subset of neurons. The encoded memory is then recalled by increasing external input to a subset of neurons within it. **G.** Histograms of mean firing rates of the assembly (orange) and the rest of the network (grey) at the four time points in **F**. The color-coded cross indicate the mean firing rates of each subset of neurons. **H.** The evolution of the firing rates of the assembly (orange) and the rest of the network (grey) during the memory task. The shaded area indicates STD over all neural subset units. 48
- 2.5 A** The external current supplied to the CA3 neurons. Different colours indicate different place fields. **B** The CA3 neuron firing rates. Different colours indicate different CA3 neurons tuned to different place fields, as in **A**. 63
- 2.6 A** Firing rate of the CA1 inhibitory neurons in the first (grey) and last (black) lap of the simulation. 63
- 2.7 A** The initial CA3 to CA1 synaptic weights of neuron 1, as in Figure 2.1C. Different colours indicate synapses from CA3 neurons with different place fields, as in S1. **B** The initial CA3 to CA1 synaptic weights of neuron 2, as in Figure 2.1C. **C** The final CA3 to CA1 synaptic weights of neuron 1, as in Figure 2.1F. **D** The final CA3 to CA1 synaptic weights of neuron 2, as in Figure 2.1F. . . . 64

- 2.8 **A** CA3 to CA1 synaptic weights of an example place (green) and silent (grey) cell as a function of lap number. **B** Mean CA3 to CA1 synaptic weights of an all place (green) and silent (grey) cell as a function of lap number. The bold lines denote the mean over all active cells. **C** CA1 inhibitory synaptic weights as a function of time. The bold lines denote the mean value per lap. **D** The synaptic input to the inhibitory neurons (grey line) as a function of time. The bold black line denotes the mean per lap. The red dotted line denotes the target input. . . . 65
- 2.9 **A** The number of active place cells as a function of target input value. The error bars denote the mean over 20 networks. 65
- 2.10 **A** Evolution of the inhibitory firing rates during the silencing protocol as described in Figure 2.2A. **B** Evolution of the inhibitory synaptic weights during the silencing protocol as described in Figure 2.2A. 66
- 2.11 **A** The evolution of all the synaptic weights from CA3 to CA1 for an example place (green) and silent (grey) cell during the consolidation protocol as described in Figure 2.2I. Dashed vertical lines mark testing laps. **A** Same as in **A** , but for the mean synaptic weights an example place (green) and silent (grey) cell. . . . 66
- 2.12 Recurrent network. **A** All the inhibitory synaptic weights. **B** All the excitatory firing rates. 67
- 2.13 **A-E** Same as in Figure 2.3B,E,F,C but for the inhibitory neurons only. . . . 67
- 2.14 **A-D** Same as in Figure 2.3B,E,F,C but for the network with fixed inhibitory-inhibitory connections. The inhibitory neurons are connected with a probability $p_{II} = 0.2$, with the in-degree kept constant. 68
- 2.15 **A-D** Same as in Figure 2.3B,E,F,C but for a bigger network ($N_E=400$, $N_I=100$). Here we use $1/N$ scaling for W_{rec} 69

- 2.16 **A** Firing rate as a function of time with varying learning rate η_{IDIP} . The shaded area denotes the standard deviation of the firing rates over all the neurons in the network. **B** Firing rate distribution of networks with different learning rates η_{IDIP} after inhibitory learning. Here we include data from 20 different networks. **C** ISI CV of networks with different learning rates after inhibitory learning. Here we include data from 20 different networks. 70
- 2.17 **A-D** Same as in Figure 2.3**B,E,F,C** but for the network following iSTDP (Vogels et al., 2011). **C** The pale orange histogram indicates the distribution of firing rates with IDIP in the same network. 71
- 2.18 **A** Standard deviation of the excitatory firing rates σ_{fr} before (x-axis) and after (y-axis) inhibitory learning for the networks following IDIP (Pearson correlation coefficient $r = 0.59$, p-value = $2.8e-22$). Different standard deviation of the W_{rec} sampling distribution (σ_{wrec}) values are color-coded. **B** Same as in **A** but for networks following iSTDP (Vogels et al., 2011) (Pearson correlation coefficient $r = 0.08$, p-value = 0.22). **C** Final firing rate distribution of networks following IDIP with different W_{rec} sampling distribution standard deviation values σ_{wrec} . Here we include data from 20 different networks. **D** Same as in **C** for ISI CV. 72
- 2.19 **A** The distribution of the firing rates after IDIP as a function of different values of inhibitory probability of connections (horizontal axis) and the target input θ_i (vertical axis). Each histogram contains data from 20 different networks. The means of each distribution are denoted with black triangles. As the target input value increases (y axis), we observe an increase in the mean firing rate (black triangles) and an increase in firing rate diversity, as the network is less constrained. As the inhibitory connectivity increases (x axis), the excitatory population subsets that each inhibitory neurons is connected to increasingly overlap, and thus there is a decrease in the diversity of the excitatory firing rates. 73
- 2.20 **A** The mean firing rate of networks with varying excitatory connectivity. The error bars denote the standard deviation over 20 networks. 74

- 2.21 **A-B** Recurrent network where a subset of neurons receives increased input, as in Figure 2.4A. . **A** The evolution of all the inhibitory synaptic weights **B** Same as Figure 2.4B but in a network following iSTDP (Vogels et al., 2011). **C-D** Recurrent network as in Figure 2.3A, with no neurons receiving increased external input. **C** The excitatory firing rate before (x-axis) versus after (y-axis) inhibitory learning in the network following IDIP. (Pearson correlation coefficient $r = 0.29$, p-value = $5.09e-32$) **D** The excitatory firing rate before (x-axis) versus after (y-axis) inhibitory learning in the network following iSTDP (Vogels et al., 2011) (Pearson correlation coefficient $r = 0.005$, p-value = 0.84). 75

- 2.22 **A** All the inhibitory synaptic weights during the associative memory task (Figure 2.4F). The dashed vertical lines indicate memory encoding and recall respectively. **B** The assembly firing rates during recall. **C-D** Same as Figure 2.4G-H but in a network following iSTDP (Vogels et al., 2011). 76

- 2.23 **A** The final firing rate rank correlation coefficient, calculated as in Figure 2.4E, for the recurrent networks following different IDIP learning rule variants (outlined in the supplementary methods). The correlation coefficient values of the networks following IDIP and iSTDP are significantly different (student t-test, p value = $7e-12$). The correlation coefficient values of the networks following IDIP and IDIP learning rule variants are not significantly different (student t-test, p values = 0.07 (v1), 0.96 (v2), 0.47 (v3), 0.09 (v4), 0.18 (v5)). **B** The mean firing rate of the place map before ('OFF') and during ('ON') silencing for the hippocampal networks following IDIP learning rule variants. Gray circles indicate individual trials. Colour-coded circles indicate average over all 20 trials. 77

- 2.24 **A** Diagram of a network with sparse inhibitory connectivity. Two excitatory (triangles) and two inhibitory (circles) neurons are simulated. One of the excitatory neurons (green) receives a slightly (5%) stronger external input. Thus one inhibitory neuron (blue) receives synaptic input from the excitatory neuron that is more active and projects onto the other excitatory neuron (purple). The other inhibitory neuron (orange) receives synaptic input from the less active excitatory neuron (purple) and projects onto the more active (orange) excitatory neuron. **B** The evolution of neuronal firing rates for the excitatory cells, color coded as in **A**. The initially small activity difference is amplified by the inhibitory plasticity, leading to increased activity of one excitatory neuron (green) and a decrease in the other (purple). **C** The synaptic input to the inhibitory neurons as a function of time, color coded as in **A**. The dashed line denotes the target input. **D** The evolution of the inhibitory (I to E) synaptic weights, color coded as in **A**. **E-H** The same as in **A-D** but without any plasticity. 78

- 3.1 **Recurrent neural networks with feedback control feature rapid adaptation to *acute* task perturbation and can learn the *persistent* task perturbation using a local learning rule.** A recurrent neural network with feedback control is first pre-trained on produce a 2D sigmoidal velocity profile based on a target and *go* cue input input, and then tested on a 30° target rotation. (a) The pre-trained network adapts to the perturbation during the first trial with feedback control (full-line). The dashed line denotes the same network without feedback control. (b) The network further improves its performance during persistent perturbation with feedback-driven local learning (Feulner et al., 2022) 90

3.2 Feedback control modulates rapid adaptation during task perturbation through approximate inference learning in the activity space.

(a) Increasing task perturbations degree ($0 \rightarrow 30 \rightarrow 60$) increases the relative feedback contribution to the network readout during a single task trial, which included a complete reaching movement. **start** denotes the start of the movement, and **peak** denotes the velocity peak. The shaded region denotes the standard error of the mean over 20 network seeds and all 8 targets. (b) The average magnitude of the feedback drive is proportional to the movement perturbation magnitude. (c) The cosine similarity of the online feedback signal injected into the ongoing network dynamics and the optimal, global trial gradient w.r.t. activations during a single task trial. The shaded region denotes the standard error of the mean over 20 network seeds and all 8 targets. 92

3.3 Feedback control aids local learning in the recurrent network layer.

(a) Train loss as a function of trial number during persistent perturbation with feedback-driven local learning. Grey line denotes the tested performance of the same network without feedback control, showing that the network relies less on feedback control with learning. The shaded regions denote the standard deviation over 10 random network seeds. (b) The final performance (after 1000 trials) of recurrent networks trained with feedback-driven local learning with (blue) or without (yellow) feedback control during learning with varying degrees of persistent perturbation. The points in red denote the performance of networks trained with online BPTT. Note that all the performances reported here are in the setting without feedback control at test-time, irrespective of if feedback control was used during training. The error bars denote the standard deviation over 10 random network seeds. 93

- 3.4 Feedback increases the accuracy of local learning.** (a) The norm of the Jacobians of the network activities at time t w.r.t. to activities at a previous timestep during a single trial. The shaded regions denote the standard deviation over 10 random network seeds. (b) The mean local alignment of the feedback-driven local learning rule with the true local gradient during task adaptation. The error bars denote the standard deviation over 10 seeds. The error bars denote the standard deviation over 10 random network seeds. (c) The mean global alignment of the feedback-driven local learning rule with the true global gradient during task adaptation. The error bars denote the standard deviation over 10 random network seeds. 94
- 3.5 Feedback increases the efficiency of local learning.** (a) The mean norm or efficiency ratio of the total, global weight update with that of the online, local weight update during task adaptation at various perturbation degrees. The error bars denote the standard deviation over 10 random network seeds. (b) The cosine similarity between the online feedback control signal injected into the ongoing network dynamics and the second-order gradient (calculated using the Newton method) w.r.t. activations (blue) during a single trial. The grey line denotes the first-order gradient as in Figure 3.2c. The shaded region denotes the standard error of the mean over 10 network seeds and all 8 targets. 96
- 3.6 Performance of networks trained using the local rule from Feulner et al. (2022) with increasing degrees of perturbation.** 101
- 3.7 Performance of networks trained using RFLO learning from Murray (2019) with increasing degrees of perturbation.** 101

4.1	A simple student-teacher setup reveals a performance gap between backpropagation and feedback alignment. A feedforward <i>student</i> network is trained to approximate the output generated by a randomly initialised and fixed <i>teacher</i> network. Both networks receive the same i.i.d. input. (a) Test loss of a (3 hidden layer) student networks as a function of (4 hidden layer) teacher network width. The student networks have 64 hidden units per layer and are trained with either backpropagation (BP), feedback alignment (FA), direct feedback alignment (dFA) or delta rule (DR). With DR, only the readout layer is plastic, and no error propagation occurs to the deeper layers. Here, the error bars denote the standard deviation over 4 randomly initialized student <i>and</i> teacher networks. (b) Test loss of the same student network as a function of number of presented examples. Here, the learning task is a teacher network with a width of 1024 hidden units.	113
4.2	Meta-learning identifies a limited set of learning rule parameters for the direct feedback weights. (a) Meta loss as a function of meta iterations. (b) Parameter α , which denotes the proportion of the error coming from regular feedback connections, as a function of meta iterations. (c) Regular feedback weight parameters as a function of meta iterations. (d) Direct feedback weight parameters as a function of meta iterations.	114
4.3	Meta-learned learning rule closes the performance gap between backpropagation and direct feedback alignment and aligns the pseudo-gradients with a limited number of examples. (a) Inner loss as a function of inner iterations for the meta-learned learning rule (meta, light blue), backpropagation (BP, dark blue) and direct feedback alignment(dFA, orange). (b) Inner gradients as a function of inner iterations for the meta-learned learning rule and direct feedback alignment. Color-coded as in (a). Shaded region indicates the standard deviation over 16 randomly initialized student networks and 2 teacher networks. (c) Per-layer 2-norm of final weight updates for the different learning rules, color-coded as in (a).	116

- 4.4 Variance of the teacher network output as a function of hidden layer size. Note the importance of the gain in weight initialization for the expressiveness of the network, as it pushes the *tanh* activation function into the non-linear regime. . 119

Chapter 1

Introduction

Consider a child learning to ride a bicycle or an adult mastering a new language: these feats are made possible by the brain's remarkable ability to adapt its neural pathways – a phenomenon known as neural plasticity. This capacity for continual adaptation and learning, encompassing everything from fine-tuned synaptic adjustments to large-scale neural network reorganization, has long captivated neuroscientists. Given the immense complexity of this phenomenon, experimental studies and computational modelling are often employed in tandem to unravel its intricacies. This thesis explores various computational approaches to understanding biological learning, beginning with a brief overview of synaptic plasticity and its associated challenges.

1.1 Computational models of biological learning

Neurons and their communication. Neurons, the brain's information processors, collect and transmit signals through the release of neurotransmitters and neuromodulators. The presence of these around the neuron, specifically around the synapses on its dendritic tree, changes the polarization of its membrane. When its membrane depolarization exceeds a threshold, a neuron fires a rapid, all-or-nothing spike (or spike train). Spikes trigger neurotransmitter release at synapses, the junctions between neurons. Synaptic transmission is generally unidirectional, with effects determined by the presynaptic neuron. Neurons can be excitatory, promoting signalling, or inhibitory, suppressing it (Dale, 1935). The precise interplay between excitatory and

inhibitory neurons is thought to enable the complex processing capabilities of neural circuits.

Modelling neurons. Neurons can be computationally modelled at different levels of abstraction. Spike-based models offer a detailed, temporally precise representation of neuronal activity, capturing the timing of individual action potentials and thus encoding information in their frequency. One common model for this is the Leaky Integrate-and-Fire (LIF) model (Gerstner and Kistler, 2002), which describes the membrane potential dynamics with a differential equation:

$$\tau_m \frac{dV(t)}{dt} = -(V(t) - V_{\text{rest}}) + RI_{\text{ext}}(t), \quad (1.1)$$

where τ_m is the membrane time constant, $V(t)$ is the membrane potential at time t , V_{rest} is the resting potential, R is the membrane resistance, and $I_{\text{ext}}(t)$ is the external input current. When $V(t)$ reaches a threshold V_{th} , the neuron fires a spike, and $V(t)$ is set to the neuron reset potential V_{reset} . In contrast, rate-based models offer a computationally simpler approach by averaging neural activity over time, trading the exact temporal precision of each neuronal spike for computational efficiency. These models are particularly useful for studying large-scale network dynamics and general activity patterns. A commonly used rate-based model is described by the equation:

$$\tau \frac{dr_i}{dt} = -r_i + \phi \left(\sum_j w_{ij} r_j + I_i \right) \quad (1.2)$$

where r_i represents the firing rate of neuron i , τ is the time constant, ϕ is a nonlinear activation function, w_{ij} are the synaptic weights, and I_i is the external input to neuron i .

These models serve different purposes: spike-based models are useful for understanding temporal dynamics and precise spike timing, while rate-based models are efficient for studying the overall behavior of neural networks. We will be using both in this thesis.

Synaptic plasticity. Synaptic signal transmission efficacy is not constant; it dynamically changes through experience and neural activity (Bliss and Lømo, 1973; Buonomano and Merzenich,

1998). These changes, collectively known as synaptic plasticity, occur through a wide range of biological mechanisms (Cooke and Bliss, 2006; Ho et al., 2011). These mechanisms include changes in the density of synaptic neurotransmitter receptors (Martin et al., 2000), modifications to intrinsic neuronal excitability (Desai et al., 1999), alterations in gene regulation (Zovkic, 2021), and the birth and death of synapses themselves (Lamprecht and LeDoux, 2004). This diverse array of physiological changes is believed to be the key mechanism through which animals adapt and learn (Ramirez and Arbuckle, 2016). Although the experimental data helps explain the processes between individual synaptic changes, it can fall short at explaining the global network effects of such local changes. This is precisely where computational modelling excels, and allows us to integrate diverse experimental findings to elucidate mechanisms of network-wise learning and adaptation.

Modelling Synaptic Plasticity. Mathematically, we can describe synaptic plasticity by learning rules that capture the complex relationship between neural activity and changes in synaptic efficacy. Unlike the learning rules in machine learning, these rules are typically constrained by local parameters, such as the activity of presynaptic and postsynaptic neurons and the current strength of the synaptic connection. When a learning rule relies solely on the statistics of neural activity to drive synaptic modifications, it is classified as *unsupervised* learning. A classic example of an unsupervised learning rule is Hebb’s postulate, which proposes that synaptic strength is potentiated when pre- and postsynaptic neurons fire in close temporal proximity (Hebb, 1962). This is often summarized as “*neurons that fire together, wire together*”. Mathematically, in rate-based neurons Hebbian learning can be expressed as:

$$\Delta w_{ij} = \eta r_i r_j \tag{1.3}$$

where Δw_{ij} is the change in synaptic weight, η is the learning rate, and r_i and r_j are the firing rates of the postsynaptic and presynaptic neurons, respectively. In contrast, supervised learning involves a feedback signal that guides the direction and magnitude of synaptic modifications to correlate with improved performance. This type of learning can be represented by a learning

rule that adjusts synaptic weights based on the difference between the actual postsynaptic firing rate and some target firing rate. A simplified version of a supervised learning rule might be expressed as:

$$\Delta w_{ij} = \eta(r_i - \hat{r}_i)r_j, \quad (1.4)$$

where Δw_{ij} is the change in synaptic weight, η is the learning rate, r_i is the actual postsynaptic firing rate, and $\hat{r}_i$ is the target firing rate.

1.2 The trouble with learning in the brain

Using computational methods, this thesis takes a stab at three different but related challenges within the broader landscape of biological learning.

The challenge of maintaining network homeostasis. While Hebbian learning can generate specific neural activity patterns, it poses a challenge in excitatory networks: positive feedback can lead to runaway excitation and pathological hyperactivity. Thus, constraints and adjustments are needed to maintain a functional firing rate (Miller and MacKay, 1994; Abbott and Nelson, 2000). It is well known that neuronal activity undergoes regulation through various homeostatic plasticity mechanisms, such as synaptic scaling (Turrigiano, 2008) and the plasticity of inhibition (Vogels et al., 2011). Although these mechanisms vastly differ in their mode of action, they often implement neuron-centered homeostasis, where the deviations from the target activity of each individual neuron are suppressed. However, some experimental studies suggest that there are additional network-centered homeostasis mechanisms where the average activity of the whole network is under control (Hirase et al., 2001; Slomowitz et al., 2015; Trouche et al., 2016). This raises a fundamental question: how do networks detect and compensate for disruptions when individual neurons only sense their own activity? Inhibitory neurons have been proposed as the likely candidate solution, mainly due to their dense interconnectivity with their local network. However, these results cannot be explained by the commonly used, neuron-centred model of inhibitory plasticity (Vogels et al., 2011), which effectively sets

a target firing rate for each excitatory neuron. Thus, such learning rules over-constrain the whole network activity. This highlights the need for a deeper understanding of network-centred mechanisms that can account for the complex interactions and global regulation observed in neural networks. Chapter 2 dives more into this problem.

The challenge of *spatial* credit assignment. A fundamental challenge in biological learning is credit assignment – determining which specific synapses contribute to a particular outcome or response. This is a particular challenge in deep feedforward networks, where the neurons are layered in such a way that the contribution of each neuron to the final output is not immediately apparent, making it difficult to trace the origin of specific errors or successes back through the network. In the closely related field of machine learning, this problem is solved by backpropagation (BP) (Rumelhart et al., 1986), an algorithm that calculates the gradient of the loss function with respect to each weight by the chain rule, efficiently distributing the error signal backwards through the network. This means that the error signals to the deeper network layers are delivered through feedback connections that exactly mirror the feedforward connections. While the brain does feature extensive feedback connections (Peters and Payne, 1993; Peters et al., 1994), it does not have the required strictly reciprocal connectivity, which makes backpropagation biologically implausible in its direct implementation (Lillicrap and Santoro, 2019). This specific challenge in biological learning has been termed the “weight transport” problem (Grossberg, 1987). Although recent work has shown that this symmetry is not strictly necessary (Lillicrap et al., 2016; Nøkland, 2016; Liao et al., 2016; Akrouit et al., 2019; Kunin et al., 2020), how biological systems could efficiently solve the credit assignment problem remains an active area of research. Chapter 4 dives more into this problem.

The challenge of *temporal* credit assignment. An additional layer of complexity in biological learning arises from the temporal dimension. Often, actions or events that occur far apart in time can contribute to a desired outcome. Identifying and reinforcing the specific synapses involved in those earlier events, even if their contribution isn’t immediately apparent, is crucial. In machine learning, this is solved by backpropagation’s temporal analogue,

backpropagation-through-time (BPTT) (Werbos, 1990). Here, a network is “unrolled” over some range of time steps in the past, and backpropagation is performed as if the network was a feedforward one, with each time step now representing a layer in this unfolded network. This allows the gradients to be calculated and propagated through time, effectively adjusting the weights based on temporal dependencies. This poses additional challenges in terms of biological plausibility. The first one is the act of “unrolling” itself, being opposite in (time) direction of forward network propagation. This can be effectively solved with the real-time recurrent learning (RTRL) algorithm (Williams and Zipser, 1989). However, RTRL is often more computationally demanding than BPTT, which poses an implementation challenge both in the brains and machines (Marschall et al., 2020). It has been shown that a significant approximation of RTRL can be both biologically plausible and effective under certain conditions, by relying on short-term dependencies present in neural activity at the point of learning (Murray, 2019; Bellec et al., 2020). Chapter 3 dives more into this problem.

1.3 Thesis structure

Chapter 2. The first part of this thesis focuses on the problem of network homeostasis in the biological networks. Specifically, a novel form of inhibitory homeostatic plasticity is examined, which allows for global control of the network rather than just local control. Moreover, this form of plasticity is functionally analyzed within the context of both the hippocampus and the cortex. The main contribution of this chapter is the integration of inhibitory plasticity, neural representation remapping, and network homeostasis into a single robust framework.

Chapter 3. The second part of this thesis focuses on the problem of temporal credit assignment. Specifically, the mechanisms behind the previously proposed rapid and biologically plausible model of motor adaptation to task perturbation are examined. The main contribution of this chapter is the finding that feedback control enables both accurate and efficient biologically plausible learning in the notoriously difficult to train recurrent neural networks.

Chapter 4 The third and final part of this thesis focuses on the problem of spatial credit assignment. Specifically, it investigates how learning in cortical multi-layer networks occurs through spatially and temporally local synaptic plasticity aided by feedback connections from upper layers. In contrast to the previous chapters, this chapter uses the meta-learning framework to explore the space of effective biologically plausible feedback learning rules. The main contributions are the proposition of teacher-student network setup as an evaluation task for effective credit assignment, and the discovery of a novel feedback learning rule.

Chapter 2

Network-centered homeostasis through inhibition

This chapter explores a novel mechanism for network-based inhibitory plasticity, offering a mechanistic explanation for the emergence of active and silent place cells in a hippocampal CA1 network model. The proposed plasticity rule explains fast and reversible remapping following place map inactivation and new place map consolidation after repeated remapping, which were previously unexplained experimental observations. Additionally, the model provides rapid firing rate homeostasis in a recurrent cortical network while preserving essential network features such as firing rate heterogeneity, neural representation, and persistent activity after memory encoding. Overall, the findings present a circuit-level understanding of the CA1 network, link various experimental data, and propose experimentally testable predictions. This chapter includes an expanded version of research first presented at COSYNE 2020 and later published in the Journal of Cell Reports.

Kaleb, Pedrosa and Clopath (2020). Balancing excitation and inhibition in spiking networks with Input Dependent Inhibitory Plasticity (IDIP). COSYNE 2020 28 February / I-16 in Denver, Colorado, USA. [\[link\]](#)

Kaleb, K., Pedrosa, V., and Clopath, C. (2021). Network-centered homeostasis through inhibition maintains hippocampal spatial map and cortical circuit function. *Cell Reports*, 36(8)

Network-centered homeostasis through inhibition maintains hippocampal spatial map and cortical circuit function

Klara Kaleb, Victor Pedrosa, Claudia Clopath

Bioengineering Department, Imperial College London, London, United Kingdom

2.1 Abstract

Despite ongoing experiential change, neural activity maintains remarkable stability. Such stability is thought to be mediated by homeostatic plasticity and is deemed to be critical for normal neural function. However, what aspect of neural activity does homeostatic plasticity conserve, and how it still maintains the flexibility necessary for learning and memory, is not fully understood. Homeostatic plasticity is often studied in the context of neuron-centered control, where the deviations from the target activity for each individual neuron are suppressed. However, experimental studies suggest that there are additional, network-centered mechanisms. These may act through the inhibitory neurons, due to their dense network connectivity. Here we use a computational framework to study a potential mechanism for such homeostasis, using experimentally inspired, input-dependent inhibitory plasticity (IDIP). In a hippocampal CA1 spiking model, we show that IDIP in combination with place tuned input can explain the formation of active and silent place cells, as well as place cells remapping following optogenetic silencing of active place cells. Furthermore, we show that IDIP can also stabilise recurrent network dynamics, as well as preserve network firing rate heterogeneity and stimulus representation. Interestingly, in an associative memory task, IDIP facilitates persistent activity after memory encoding, in line with some experimental data. Hence, the establishment of global network balance with IDIP has diverse functional implications and may be able to explain experimental phenomena across different brain areas. Finally, we make testable experimental predictions.

2.2 Introduction

Although neural activity varies, presumably due to different demands of various cognitive functions, it is usually constrained to a small operational range of only a few Hertz. This regime may facilitate energy saving (Perrinet, 2010), as well as optimal information processing (Laughlin, 1981; Stemmler and Koch, 1999). Furthermore, deviations from this range are often associated with pathological states, such as epilepsy, schizophrenia, and Fragile X syndrome (Dickman and Wondolowski, 2013). Given that neural activity undergoes constant experiential change, there exists a requirement for active processes to maintain stability. Many such processes have been identified, such as synaptic scaling (Turrigiano et al., 1998; Desai et al., 2002; Turrigiano and Nelson, 2004; Goel and Lee, 2007; Glazewski et al., 2017), intrinsic plasticity (Desai et al., 1999; Gainey et al., 2018; Lambo and Turrigiano, 2013; Maffei and Turrigiano, 2008), meta-plasticity (Bienenstock et al., 1982; Kirkwood et al., 1996; Zenke et al., 2013; Frank et al., 2006), diffusive neuromodulation (Sweeney et al., 2015; Steinert et al., 2008, 2011; Naumann and Sprekeler, 2020), structural plasticity (Yin and Yuan, 2015; Gallinaro and Rotter, 2018) and inhibitory plasticity (Woodin et al., 2003; Maffei et al., 2004, 2006; Chen et al., 2011; Vogels et al., 2011; Keck et al., 2011; van Versendaal et al., 2012; D’amour and Froemke, 2015; Udakis et al., 2019; Clopath et al., 2016; Hennequin et al., 2017; Das et al., 2011; Haas et al., 2006). These altogether make up the term homeostatic plasticity and, although vastly different, they all act as a negative feedback mechanism that adjust the neural parameters to compensate for deviations from some set-point. Homeostatic plasticity is often studied in the context of neuron-centered control, as neurons return to their preferred level of activity after experimental manipulations that decrease (Hengen et al., 2013) or increase activity (Pacheco et al., 2019). There is also experimental evidence of network-centered homeostasis (Hirase et al., 2001; Slomowitz et al., 2015; Trouche et al., 2016), where instead of the individual neurons, the mean activity of the whole network is homeostatically maintained. However, computational studies of such mechanisms are few (Sweeney et al., 2015; Mackwood, 2019; Naumann and Sprekeler, 2020) and thus they remain less well understood.

To illustrate network-centered homeostasis, we turn to the hippocampus. Spatial environments

are known to be represented by a cognitive map, consisting of a subset of hippocampal pyramidal cells. These are called place cells, as they fire action potentials when the animal is in a specific location within the environment, their place fields (O'Keefe and Dostrovsky, 1971; O'Keefe, 1976; O'Keefe and Nadel, 1978; Wilson and McNaughton, 1993). Optogenetic silencing of the CA1 pyramidal neurons encoding a place map of a familiar environment leads to rapid activation of previously silent CA1 pyramidal neurons (Trouche et al., 2016). This abrupt increase is followed by a slower, seconds-long activity change towards a stable level on par with that of the original place map (Trouche et al., 2016). Thus, an alternative place map transiently emerges while the original place cells are being silenced, and the spatial representation is homeostatically maintained. Furthermore, with repeated silencing, the alternative place map is consolidated over the original place map (Trouche et al., 2016). As neurons presumably have access only to their own activity, it is unclear how such a perturbation can be detected and compensated for at the network level.

One of the likely candidates to implement such network-centered homeostasis are the inhibitory neurons. CA1 inhibitory neurons are strongly connected to a large number of heterogeneously tuned CA1 pyramidal cells (Ali et al., 1998; Freund and Buzsaki, 1996; Gulyás et al., 1999; Bezaire and Soltesz, 2013; Csicsvari et al., 1998a; English et al., 2017; Csicsvari et al., 1998b), and exhibit broad spatial tuning (Grienberger et al., 2017). Thus, they can facilitate lateral inhibition between the pyramidal cells (Geisler et al., 2007) and as such, both sense and influence the activity of their local network. Hence, plasticity in the inhibitory circuitry could offer a highly efficient network activity homeostasis, which has previously been shown with neuron-centered plasticity (Vogels et al., 2011). However, there are indications that inhibitory plasticity can also act more globally. For instance, inhibitory postsynaptic potentials (iPSPs) have been shown to decrease when global, but not single neuron, spiking activity is suppressed (Hartman et al., 2006; Peng et al., 2010). Moreover, in the highly recurrent pyramidal cell networks of the neocortex, where inhibitory neurons are similarly interconnected with the rest of the network (Sohya et al., 2007; Niell and Stryker, 2008; Kerlin et al., 2010; Ma et al., 2010; Zariwala et al., 2011; Znamenskiy et al., 2018; Wilson et al., 2017; Fino and Yuste, 2011; Packer and Yuste, 2011; Hofer et al., 2011; Bock et al., 2011; Pfeffer et al., 2013), inhibitory scaling was

shown to be Ca^{2+} /calmodulin-dependent protein kinase IV (CAMKIV) independent (Joseph and Turrigiano, 2017), and thus may be decoupled from unique postsynaptic neuron activity. Furthermore, sensory deprivation studies across the primary cortices (Kuhlman et al., 2013; Barnes et al., 2015; Li et al., 2014; Gainey et al., 2018) have led to suggestions that inhibition could be broadly adjusted as a function of the activity of the circuit (Gainey and Feldman, 2017). Inspired by all of these experimental findings, we hypothesize that the synaptic input to the inhibitory neurons could act as a proxy for the local network activity, and as such be used to adjust the level of inhibition appropriately. Such plasticity could occur through depolarization induced modulation of spike-evoked inhibitory transmission (Christie et al., 2011; Bouhours et al., 2011; Rowan and Christie, 2017) or inhibitory neuron intrinsic excitability (Gainey et al., 2018).

In this work, we computationally study the properties and potential functionality of such novel inhibitory plasticity, which we term input-dependent inhibitory plasticity (IDIP). First, we show that such plasticity provides a mechanistic explanation for the emergence of active and silent place cells in a hippocampal CA1 network model. Our model also reproduces the previously unexplained experimental data of fast and reversible remapping following acute optogenetic place map silencing (Trouche et al., 2016), as well as the consolidation of the alternative place map following repeated silencing (Trouche et al., 2016). Furthermore, we show that IDIP in a cortical recurrent network model provides a form of rapid firing rate homeostasis while maintaining important network features, such as firing rate heterogeneity and persistent activity. Thus, we show that IDIP allows for accurate maintenance of neural representation while preserving flexibility important for neural coding.

2.3 Results

2.3.1 Input-dependent inhibitory plasticity (IDIP) rule as a homeostatic mechanism

Experimental studies suggest an existence of network-centered homeostasis (Hirase et al., 2001; Slomowitz et al., 2015; Trouche et al., 2016), where the mean firing rate of the network, rather than that of the individual neurons, is homeostatically maintained. For example, rapid homeostasis of place representation in the hippocampal CA1 region is observed after optogenetic silencing of a familiar place map, through the emergence of an alternative place map (Trouche et al., 2016). The likely candidates underlying such a mechanism are the inhibitory neurons, due to their dense inter-connectivity with the surrounding pyramidal cells, both in the hippocampus (Ali et al., 1998; Freund and Buzsaki, 1996; Gulyás et al., 1999; Bezaire and Soltesz, 2013; Csicsvari et al., 1998a; English et al., 2017; Csicsvari et al., 1998b) and in the neocortex (Sohya et al., 2007; Niell and Stryker, 2008; Kerlin et al., 2010; Ma et al., 2010; Zariwala et al., 2011; Znamenskiy et al., 2018; Wilson et al., 2017; Fino and Yuste, 2011; Packer and Yuste, 2011; Hofer et al., 2011; Bock et al., 2011; Pfeffer et al., 2013). Furthermore, rapid disinhibition has been shown to be the first step in circuit reorganization after experimental deprivation across the sensory cortices (Kuhlman et al., 2013; Barnes et al., 2015; Li et al., 2014; Gainey et al., 2018), leading to suggestions that inhibition could be broadly adjusted based on the activity of the circuit (Gainey and Feldman, 2017). Such global regulation of network activity could be achieved through the modulation of inhibition as a function of the synaptic input the inhibitory neurons receive. To this end, we hypothesize a plasticity mechanism in which strong inputs onto the inhibitory neurons lead to strengthening of the inhibitory output, whereas weak inputs onto the inhibitory neurons lead to weakening of the inhibitory output. For simplicity, we choose to implement this input-dependent inhibitory plasticity (IDIP) rule by scaling the inhibitory synaptic weights as a function of the difference between the synaptic input that each inhibitory neuron receives and a set target input value (Figure 2.1A). Note that such cell-autonomous plasticity could also be implemented through the plasticity of excitability

of the inhibitory neurons, which is not explored in this work.

2.3.2 The IDIP rule allows for the emergence of active and silent place cells in a hippocampal network

To assess whether IDIP can lead to the emergence of active and silent place cells, we build a hippocampal network model of leaky integrate-and-fire neurons. The network model consists of the CA3 and CA1 region (Figure 2.1B). Each pyramidal cell in the CA3 region receives unique place tuned external current (Figure 2.5 A-B), representing the location of a simulated mouse on a 1D annular track with equally spaced place fields. The CA3 neurons project to the CA1 excitatory neurons, which are divided into equally sized groups, with neurons in each group tuned to the same place field. However, our simple model does not capture heterogeneity of tuning present in a biological system. Furthermore, as the CA1 excitatory neurons are poorly recurrently connected (Deuchars and Thomson, 1996; Thomson and Radpour, 1991), in our simulations we assume that there are no recurrent connections between them. The CA1 excitatory neurons project to the CA1 inhibitory neurons and the CA1 inhibitory neurons project back to the CA1 excitatory neurons. The CA1 inhibitory neurons have no spatial tuning (Figure 2.6A), in agreement with experiments (Dupret et al., 2013; Grienberger et al., 2017). After the first lap on the track, the excitatory synapses between the CA3 and CA1 excitatory neurons and the inhibitory synapses to the CA1 excitatory neurons are made plastic with Hebbian plasticity and the IDIP rule respectively.

We first wanted to test whether IDIP could allow for the emergence of both active and silent place cells in the CA1 region. The experiments suggest that the cells which later become active or silent are differentiable even before the first exploration of the environment (Epsztein et al., 2011). We incorporate this in our model by introducing variability in the amplitude of the place tuning of the CA1 excitatory neurons belonging to the same group (Figure 2.1C-D, Figure 2.7A-B). As the initial inhibitory weights are set to low values, all the CA1 excitatory neurons are active and covering the whole environment in the first lap on the track (Figure 2.1E). Thus, the synaptic input received by the CA1 inhibitory neurons is high, and above the target input

(Figure 2.8D). We then turn the Hebbian plasticity and our IDIP rule on and let our simulated mouse complete 100 laps on the track, which we term the exploration phase. During this phase, the CA1 excitatory neurons increase their tuning and activity, in a positive feedback loop characteristic to Hebbian learning (Figure2.1F-G, Figure2.7.C-D). This further increases the synaptic input to the CA1 inhibitory neuron, and hence increases the inhibitory synaptic weights through IDIP (Figure2.8C) to keep the synaptic input to the inhibitory neurons at a target level (Figure2.8D). As the CA1 excitatory neurons with lower place tuning initially are less active, they are unable to escape the increasing lateral inhibition. Thus stable active and silent place cells form within each CA1 place-tuned group (Figure2.1H, Figure2.8A-B). The number of place cells that are active is a function of the target input to the CA1 inhibitory neurons (Figure 2.9A). The absence of the target firing rate for the individual CA1 excitatory neurons is essential for the observed network dynamics. Hence, we show that the IDIP rule together with Hebbian plasticity and place tuned inputs can explain the formation of active and silent place cells in a hippocampal network model.

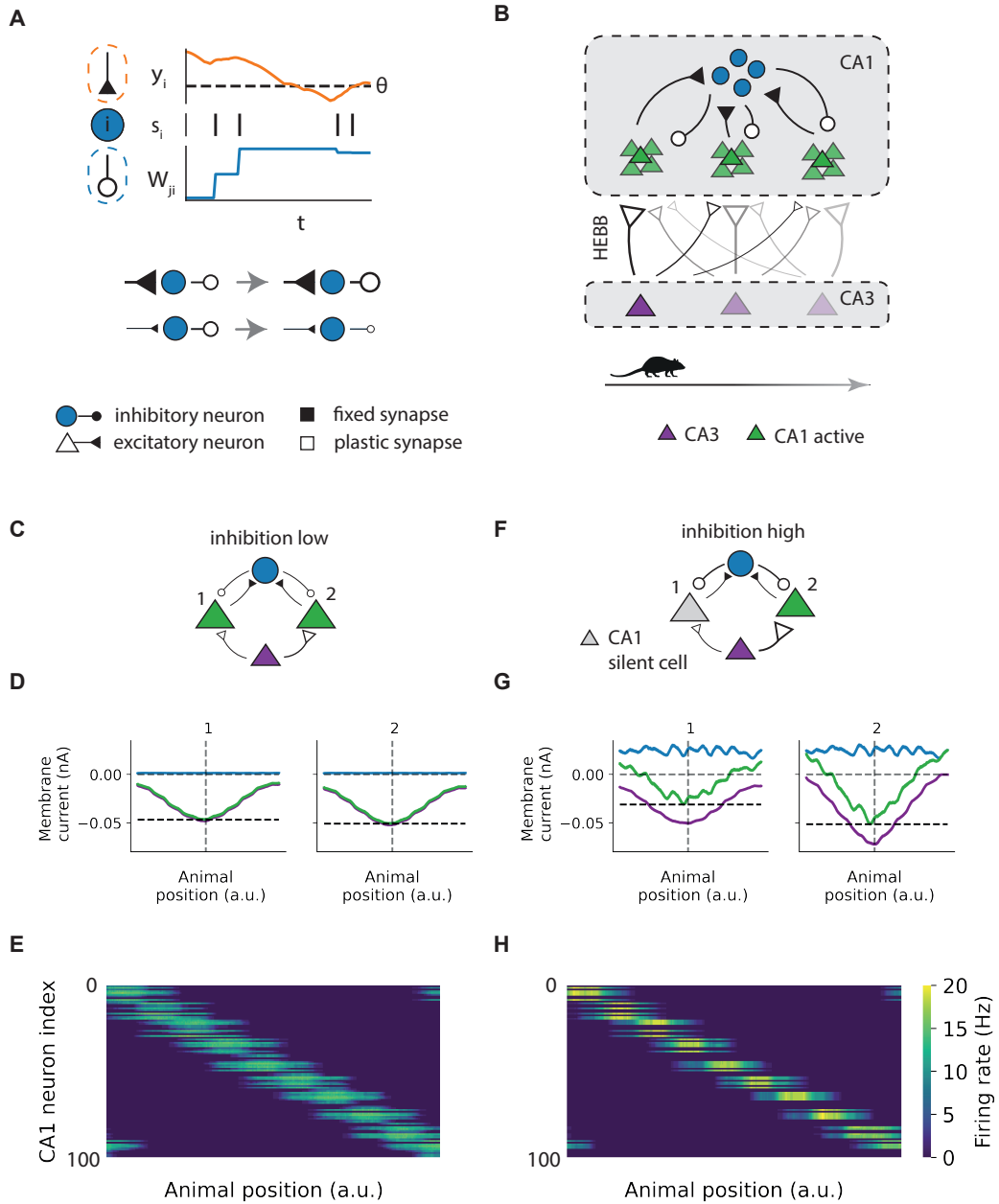


Figure 2.1: The IDIP rule enables emergence of active and silent place cells in a hippocampal network. **A.** Inhibitory plasticity rule diagram. An inhibitory neuron (blue) adjusts its inhibitory synaptic weights based on the synaptic input it receives and its target input θ . **B.** Hippocampal network diagram. The CA3 excitatory neurons (purple) receive place-dependent external current. CA3 neurons project to uniquely tuned CA1 excitatory neurons (green), which are interconnected with CA1 inhibitory neurons (blue). The CA3 to CA1 excitatory synapses are plastic under Hebbian plasticity and the CA1 inhibitory synapses are plastic under IDIP. **C-E** State of the network during the first lap, before any learning. **C.** A diagram of a sample microcircuit found in the network in **B**, consisting of 2 CA1 place cells (1 and 2) with the same spatial tuning, but with cell 2 having stronger place tuning. **D.** The excitatory (purple), inhibitory (blue) and net (green) currents received by cells 1 and 2 during the first lap on the track. The higher peak net current (black dashed line) received by cell 2 reflects its stronger tuning. **E.** The firing rates of all the CA1 place cells (y-axis) during the first lap (x-axis). All the place cells are active but with varying amplitudes, reflecting differences in their place tuning strengths. **F-H** The same as in **C-E** but after learning. Due to increased inhibition, silent (grey triangles) place cells emerge, as the difference in the peak net currents from the first lap is amplified with learning.

2.3.3 The IDIP rule enables rapid homeostatic remapping during active place cells silencing

To assess whether the proposed IDIP rule can facilitate rapid place cell remapping, we perform a silencing protocol as in Trouche et al. (2016) (Figure 2.2A). After the exploration phase, we tag all the established active place cells and silence them. In the networks with (Figure 2.2B) and without (Figure 2.2F) IDIP, such silencing leads to a decrease in the excitatory drive to the CA1 inhibitory neurons and subsequently a rapid decrease in the inhibition to the previously silent place cells, which then form an alternative place map. However, the degree of the alternative place map activation will depend on inhibitory plasticity. In the networks with IDIP during the silencing (Figure 2.2B), the dynamics of the alternative place map activation are similar to that observed experimentally (Fig 2.C) (Trouche et al., 2016). The firing rates of the alternative place cells are first increased almost immediately after silencing onset (fast phase) and are further increased within the next 1-2 seconds (slow phase) (Figure 2.2D). Our model allows us to decompose the two phases. The fast phase is due to the rapid decrease in CA1 inhibitory neuron firing rate (Figure 2.10A), as the originally active place cells, which were driving their activity, are no longer active. Thus the alternative place map emerges, but at a lower firing rate than the original place map, due to the place cells forming the alternative place map being less sharply place-tuned than the place cells that formed the original place map. In the networks without IDIP during the silencing (Figure 2.2F), the activity of the alternative place map does not progress beyond this phase (Figure 2.2G-H). However, in the networks with IDIP during the silencing, a second, slower phase of alternative map activation occurs as the level of inhibition gets adjusted to the lower activity and thus lower synaptic input from the alternative place cells (Figure 2.10B). Therefore the final activity of the alternative place map matches that of the original place map (Figure 2.2E). Hence, the IDIP rule acts as a homeostatic mechanism to maintain network activity during acute silencing.

When we turn silencing off, the original place map re-emerges (Figure 2.2C), in agreement with experiments (Trouche et al., 2016). However, repeating the silencing protocol (Figure 2.2I) consolidates the alternative place map (Figure 2.2J), also in agreement with experiments

(Trouche et al., 2016). In our model this happens through the gradual activity-dependent Hebbian plasticity in the CA3 to CA1 excitatory synapses (Figure 2.11A). Therefore, our model suggests that synaptic plasticity of feedforward inputs onto the CA1 pyramidal cells is a good candidate for the mechanism underlying place map consolidation.

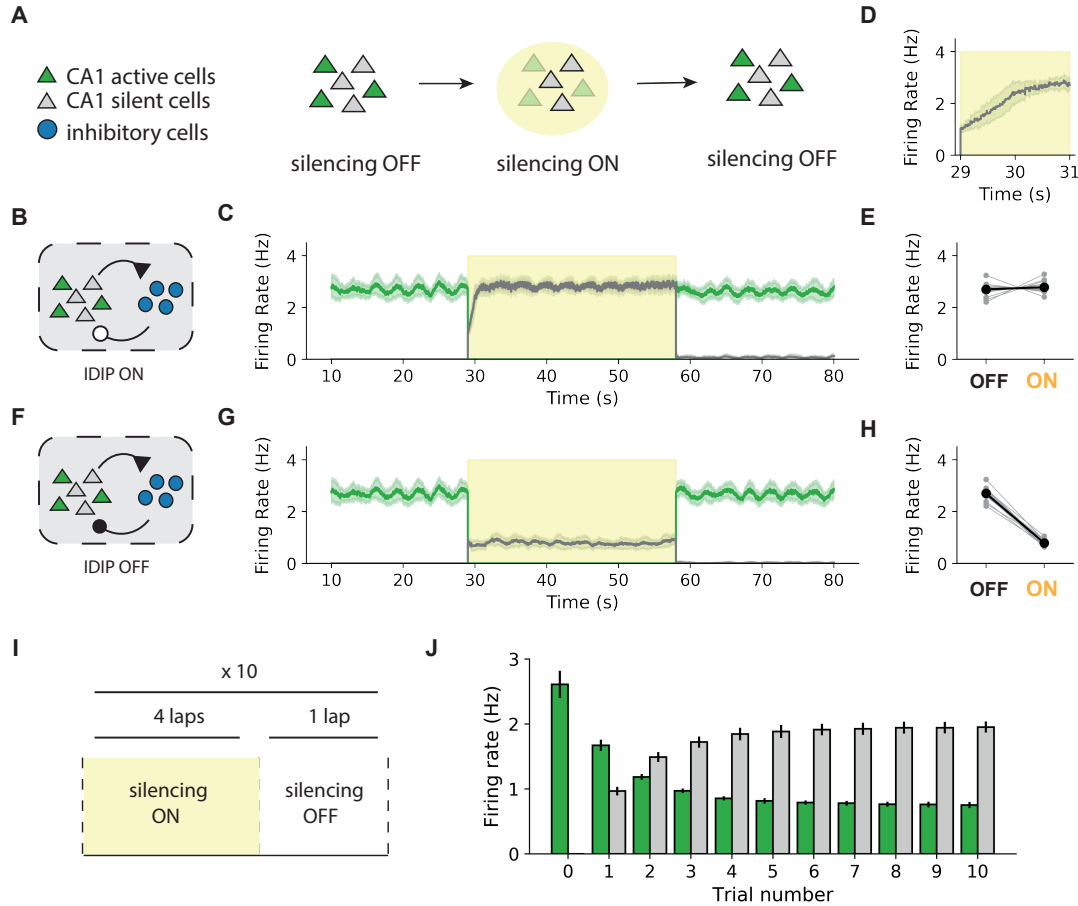


Figure 2.2: Place cells silencing leads to a rapid emergence of an alternative place map.

A. Silencing protocol similar to Trouche et al. (2016). All the active cells (green) are tagged, silenced for one lap (yellow background) and released in the subsequent lap. **B-E** Silencing in the network following IDIP. **B.** Network diagram after the exploration phase (100 laps). **C.** Mean firing rate of the network just before, during (yellow background), and after silencing. During silencing there is rapid activation of an alternative place map, consisting of the previously silent cells (grey). After silencing, the original map reactivates (green). **D.** Alternative place map activation just after silencing. **E.** The mean firing rate of the place map before ('OFF') and during ('ON') silencing. Gray circles indicate individual trials. Black circles indicate average over all 20 trials. **F-H** Same as in **B-E** but with inhibitory plasticity turned off during the silencing. **F.** Network diagram after inhibitory and excitatory learning (lap 100) with IDIP turned off during silencing (2). **G.** Without IDIP during the silencing (2) there is low alternative map activation. **H.** Without IDIP during the silencing the mean firing rate of the place map decreases. Gray circles indicate individual trials. Black circles indicate average over all 20 trials. **I** Consolidation protocol similar to Trouche et al. (2016). The established place map is silenced for 4 consecutive laps, and then released for a single testing lap. This is repeated for 10 trials. **J.** Mean firing rate of the active (green) and silent (grey) place cells during the testing laps at each trial. Trial 0 corresponds to the activity of the network just after the exploration phase.

2.3.4 The IDIP rule establishes global E/I balance in recurrent networks

The inhibitory neurons also feature dense connectivity in the neocortical circuits (Sohya et al., 2007; Niell and Stryker, 2008; Kerlin et al., 2010; Ma et al., 2010; Zariwala et al., 2011; Znamenskiy et al., 2018; Wilson et al., 2017; Fino and Yuste, 2011; Packer and Yuste, 2011; Hofer et al., 2011; Bock et al., 2011; Pfeffer et al., 2013) and there is some evidence for a network-centered control of cortical activity through inhibition (Joseph and Turrigiano, 2017; Gainey and Feldman, 2017; Gainey et al., 2018). To assess whether the proposed IDIP learning rule can homeostatically regulate such circuits, we simulate a cortical network of leaky integrate-and-fire neurons with sparse recurrent connectivity (Figure 2.3A). Each neuron in the network receives large external excitatory input and the initial inhibitory synaptic weights are set to low values. Thus, without any plasticity, the network exhibits pathologically high activity (Figure 2.3B). Due to the high activity of the network, the inhibitory neurons initially receive very large excitatory synaptic input, and thus the IDIP rule increases the inhibitory synaptic weights (Figure 2.3B - IDIP turned on at 15s, Figure 2.12A). The network, therefore, progresses from a high synchronous to a low asynchronous firing following the IDIP rule (Figure 2.3E). The excitatory firing rates reach a more physiological regime (Fig 3.B, Figure 2.12B), with a reasonable firing rate distribution and irregular firing rate dynamics (Figure 2.3C). Inhibitory neurons in the network also exhibit a diversity of firing rates and irregular firing rate dynamics (Figure 2.13A-E). Adding inhibitory-inhibitory connections to the network, increasing the network size or increasing the learning rate does not change the results (Figure 2.14, 2.15, 2.16). Hence, IDIP can homeostatically regulate network-wide activity.

The IDIP rule proposed does not impose a unique target firing rate for each neuron in the network. Instead, it controls the mean firing rate across the whole network. The final variability in the net current received by each individual excitatory neuron after inhibitory learning (Fig 3.F) results in a diversity of neuronal firing rates (Figure 2.3C). This is in contrast with the network following a neuron-centered rule, such as inhibitory spike-timing dependent plasticity (iSTDP) (Vogels et al., 2011), which effectively sets a target firing rate for the excitatory

neurons (Figure 2.17A-E, see also Figure 2.18). To assess the range of the firing rates the IDIP rule can support in a recurrent network, we simulate the network for various values of the inhibitory target input. Higher values of the inhibitory target input lead to higher network activity (Figure 2.3D, y-axis). Hence, the activity of the network following the IDIP rule depends on the target input of the inhibitory neurons. Moreover, higher inhibitory connectivity decreases mean network activity (Figure 2.3D, x-axis). As we increase the in- and out-degree of each inhibitory neuron, we observe an increase in the overlap between the populations each inhibitory neuron samples from and projects to. Thus, the dynamics of networks following IDIP are more constrained in networks with high inhibitory connectivity (see also Figure 2.19). We also show that the excitatory connectivity has no effect on the mean or variance of the network firing rate after learning (Figure 2.20). Altogether, this means that IDIP establishes a global, rather than detailed, network excitatory (E)/inhibitory (I) balance while allowing for network diversity.

2.3.5 The IDIP rule enables maintenance of neural representations

We wanted to assess whether neural representation could be maintained in the networks following the IDIP rule. To assess whether individual neurons can remain highly active, we increase the external drive to a subset of neurons in the network (Figure 2.4A). The selected subset sustains a higher firing rate relative to the rest of the network following inhibitory learning with IDIP (Figure 2.4B, S17.A). This effect is mediated by the highly active neural subset monopolizing the input to the inhibitory neurons, leading to greater inhibition to the rest of the network. Thus the rest of the network compensates for this increased activity, and the mean network firing rate is maintained (Figure 2.4C). As expected, the deviations from the target excitatory neuron firing rate are suppressed in the network following iSTDP (Vogels et al., 2011) (Figure 2.21B). Hence, IDIP can control the recurrent network activity while preserving activity heterogeneity.

To assess whether the relative firing rates between neurons can be conserved across the whole neural population, we rank the firing rates of the neurons at the beginning of our simulations,

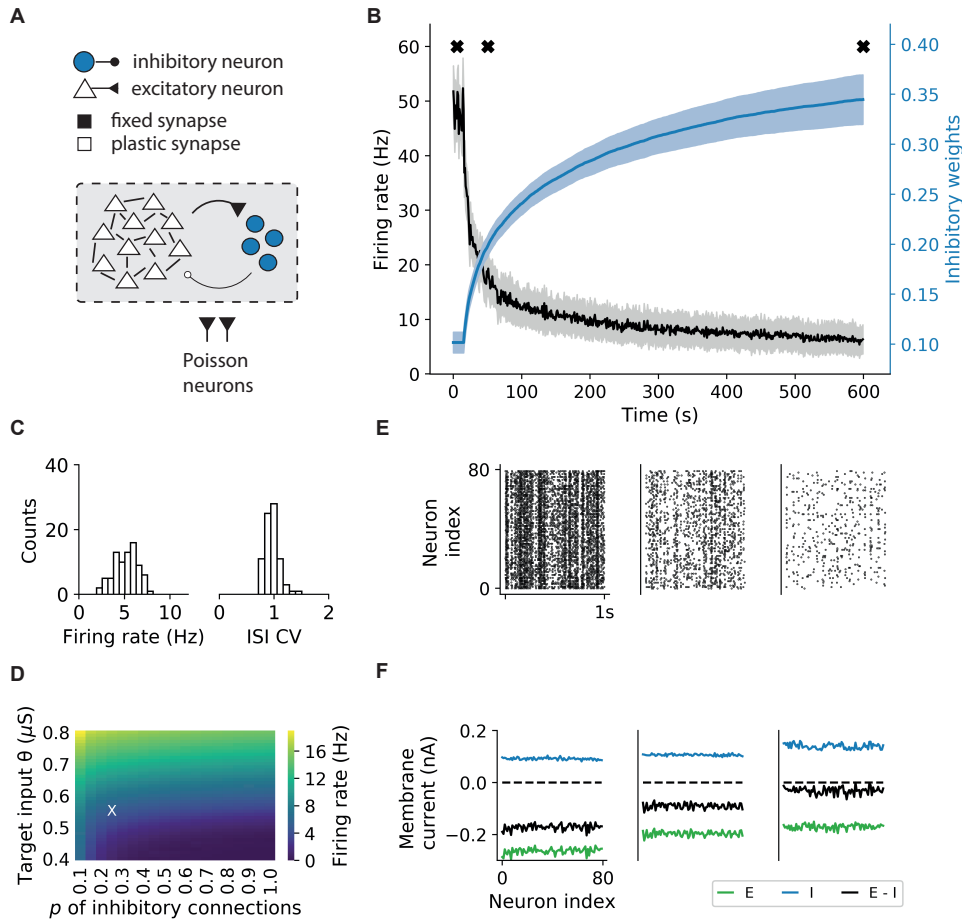


Figure 2.3: IDIP establishes global E/I balance in recurrent networks. **A.** A recurrent network of 80 excitatory (white triangles) and 20 inhibitory (blue circles) leaky integrate-and-fire neurons receiving external input from a pool of 100 Poisson excitatory neurons. Only the inhibitory synaptic weights are made plastic with IDIP. **B.** The evolution of the mean excitatory firing rate (black) and the mean inhibitory weight (blue) of the network. The shaded regions correspond to the standard deviation of individual units. IDIP is turned on at 15s. Activity of the network decreases as inhibitory weights increase. **C.** The excitatory firing rate stabilises to a mean value of 5.2 Hz (left). The network displays an asynchronous firing pattern (right). **D.** The range of mean network firing rates IDIP can support, with varying input target value (y-axis) and inhibitory connectivity (x-axis). 'X' marks the parameter combination used in our simulations. **E.** Spike raster plots at the three time points ('X') marked in **B**. The network progresses from high synchronous to low asynchronous firing following the IDIP rule. **F.** Mean membrane currents as a function of excitatory neuron index at three the time points ('X') marked in **B**.

before any inhibitory plasticity (Figure 2.4D). We then use Spearman's rank correlation coefficient to measure whether this rank is maintained over time following inhibitory plasticity (Figure 2.4D, for details see Methods). In networks following IDIP, the firing rate rank correlation is largely conserved after inhibitory learning (Figure 2.4E, Figure 2.21C). We compare this with the network following iSTDP (Vogels et al., 2011), where the firing rate rank correlation

is mostly lost, as all neurons fire at similar firing rates at the end of the simulations (Figure 2.4E, Figure 2.21D). Hence, IDIP can preserve heterogeneity in firing rates, both at a single neuron level and across the whole network.

2.3.6 The IDIP rule enables memory trace persistence and recall

We also assess the performance of the network following IDIP in a simple associative memory task. To this end, we initiate the recurrent network as before, but after inhibitory learning, we encode a memory in our network by increasing the recurrent excitatory connections within a subset of neurons (Figure 2.4F). We then attempt to recall the memory by increasing external input to a subset of the neurons belonging to the memory (for details, see methods). After encoding, the subset of neurons belonging to the memory exhibit sustained activity which is higher than the rest of the network, even after IDIP has converged (Figure 2.4G-H, Figure 2.22.A). This is in contrast to the networks following iSTDP (Vogels et al., 2011), in which the activity of the memory ensemble becomes indistinguishable from the rest of the network at convergence (Figure 2.22C-D). Such persistent activity of the memory cells has been reported in some experiments (Yassin et al. (2010); Ghandour et al. (2019), but see Barron et al. (2017)).

We then test whether we can recall the pattern, given that IDIP made inhibition stronger in the network. By increasing the external current to some units in the pattern, the memory can be recalled (Figure 2.4G-H, Figure 2.22.B). Performing the same protocol in a network following iSTDP (Vogels et al., 2011) shows that a lower number of memory cells are re-activated with recall (Figure 2.22C). Thus recall in the network following IDIP has higher fidelity. This is in line with experiments (Pignatelli et al., 2019), where increased activity of the memory cells facilitates greater pattern completion. Hence, the IDIP rule allows for sustained activity of memory cells after encoding, as well as faithful memory recall from a partial cue. Thus, we show that our proposed inhibitory plasticity rule, IDIP, can homeostatically regulate activity in models of hippocampal and cortical networks. Importantly, using IDIP we are able to reproduce previously unexplained experimental findings in the hippocampus (Trouche et al., 2016) and propose its potential functional implications in the recurrent cortical networks.

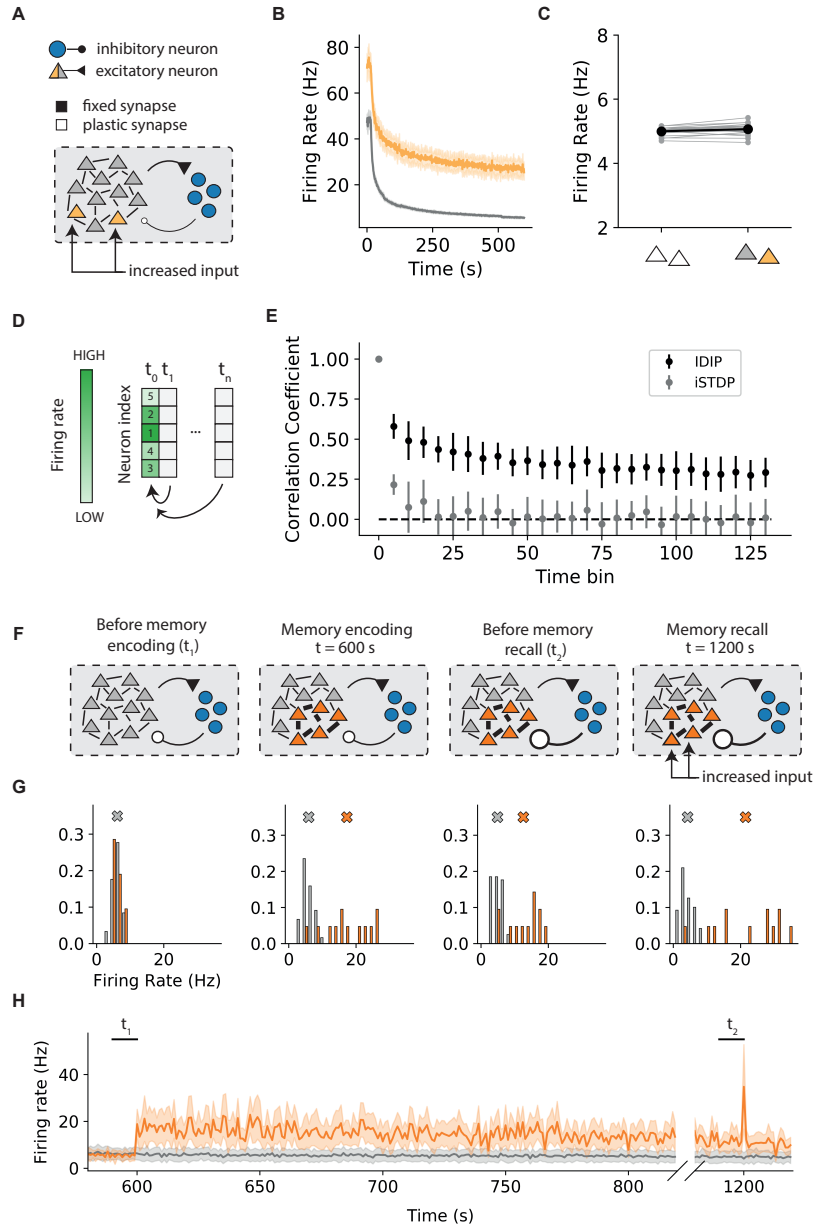


Figure 2.4: IDIP rule enables maintenance of neural representations and memory trace persistence. **A.** The same network as in Figure 2.3, but with a subset of excitatory neurons (orange) receiving increased external input (grey). **B.** The mean firing rates of the network units, color coded as in A. Bold lines denote the mean and the shaded area indicate STD over 20 trials. **C.** The mean firing rate of the networks without (left) and with (right) increased input to a subset of neurons. Gray circles indicate individual trials and the black circles indicate average over all 20 trials. **D.** Network representation task diagram. We calculate the preservation of the correlation with the initial firing rate rank during the entire course of the simulation. **E.** The Spearman rank correlation coefficient for the network following IDIP (black) and iSTDP (grey). Error bars correspond to STD over 20 trials. **F-H.** Performance of IDIP in an associative memory task. **F.** Associative memory task protocol. Following network stabilization, we increase recurrent excitatory connections between a subset of neurons. The encoded memory is then recalled by increasing external input to a subset of neurons within it. **G.** Histograms of mean firing rates of the assembly (orange) and the rest of the network (grey) at the four time points in **F**. The color-coded cross indicate the mean firing rates of each subset of neurons. **H.** The evolution of the firing rates of the assembly (orange) and the rest of the network (grey) during the memory task. The shaded area indicates STD over all neural subset units.

2.4 Discussion

Although the study of neural homeostasis is frequently neuron-centered, there is some evidence that network-centered mechanisms are also at play (Hirase et al., 2001; Slomowitz et al., 2015; Trouche et al., 2016). One of the candidates proposed for the implementation of such homeostasis are the inhibitory neurons, due to their dense connectivity, both in the hippocampus (Ali et al., 1998; Freund and Buzsaki, 1996; Gulyás et al., 1999; Bezairé and Soltesz, 2013; Csicsvari et al., 1998a; English et al., 2017; Csicsvari et al., 1998b), as well as in the neocortex (Sohya et al., 2007; Niell and Stryker, 2008; Kerlin et al., 2010; Ma et al., 2010; Zariwala et al., 2011; Znamenskiy et al., 2018; Wilson et al., 2017; Fino and Yuste, 2011; Packer and Yuste, 2011; Hofer et al., 2011; Bock et al., 2011; Pfeffer et al., 2013). In this work, we take inspiration from the experimental data and hypothesize that network homeostasis could be achieved through input-dependent inhibitory plasticity (IDIP), in which the inhibition is adjusted as a function of the synaptic input the inhibitory neurons receive. We show that in a hippocampal CA1 network model, IDIP can provide a mechanistic circuit understanding and reproduce experimental data in which place cells are optogenetically silenced, unmasking a previously silent place map (Trouche et al., 2016). Furthermore, we show that such a learning rule can also homeostatically regulate the activity of a recurrent neural network, while preserving the flexibility important for neural coding. Altogether, our results suggest that network activity homeostasis following external manipulation or endogenous changes could share a common underlying mechanism.

In contrast to the commonly used neuron-centered inhibitory plasticity (Vogels et al., 2011), the IDIP rule features an absence of a target firing rate for each excitatory neuron. Instead, with the IDIP rule, the level of inhibition is modulated to maintain the mean firing rate of the whole network. The importance of this is seen in our data-driven model of the CA1 network, where IDIP allows for emergence of both active and silent place cells. In our model, these emerge as the neurons with higher activity dominate the input to the inhibitory neurons, and thus the recruitment of lateral inhibition. Conversely, the less active cells become silent. Thus, there is an amplification of slight differences in activity through scaling of the tonic inhibition, indiscriminately to all cells. Such competition for the recruitment of inhibition has previously

been suggested to shape hippocampal assemblies (Buzsáki, 2010). Furthermore, silencing of an established place map induces rapid compensatory adjustment of inhibition, and thus an emergence of an alternative place map (Figure 2.2C), as reported experimentally (Trouche et al., 2016). We show that non-plastic inhibition is not consistent with the experimental data (Figure 2.2G). Finally, we show that repeated silencing leads to alternative place map consolidation (Figure 2.2J), as reported experimentally (Trouche et al., 2016). In our model, this occurs through gradual changes in the CA3 to CA1 excitatory synaptic weights (Figure 2.11), facilitated by disinhibition during each silencing lap (Figure 2.10A). Thus, through time, the original place map gets destabilised, which has previously been suggested as a necessary condition for remapping (Schoenenberger et al., 2016). Hence, network-centered homeostasis through IDIP provides a possible explanation for the previously unexplained experimental phenomena in the hippocampus.

We also assess whether IDIP can homeostatically regulate dynamics in the recurrent networks typical of the neocortex. Due to the absence of a firing rate set-point in the networks following IDIP, there is diversity in the neuronal firing rates within the recurrent network after inhibitory learning (Figure 2.3C). The resulting firing rate distribution is consistent with the broad and heavy-tailed distribution of firing rates observed experimentally (Wohrer et al., 2013), with some neurons more active than others (Mizuseki and Buzsáki, 2013). Such range of activity is thought to be optimal for the information storage in the brain (Laughlin, 1981) and enables linear network responses over a broad range of inputs (Sweeney et al., 2015). In our networks, highly active neurons dominate synaptic input to the inhibitory neurons, recruit inhibition more and thus inhibit other neurons more. Hence with the IDIP rule, highly active neurons can remain so as long as the rest of the network can compensate for it, providing a more flexible control. We illustrate this when we increase the external inputs to some neurons and they remain consistently more active than the rest of the network (Figure 2.4B). Furthermore, we show the IDIP rule can also extend this stability of the firing rate rank across the whole neural population (Figure 2.4E). This is not the case with inhibitory spike-timing dependent plasticity (iSTDP) (Vogels et al., 2011), as the ranks get shuffled following inhibitory learning, due to the unique firing rate set-point constraint (Fig 4.E). Hence, the IDIP model can maintain the population

firing rate around a target value while preserving neural representation. Importantly, the neural activity diversity is not imposed but emerges due to random network structure. Although such population heterogeneity can be imposed, this has been previously shown to lead to limited network responsiveness (Sweeney et al., 2015). Hence, IDIP provides a global control of recurrent network dynamics, without a firing rate set-point for each neuron, which allows for flexibility and thus conservation of network representation.

Additionally, in an associative memory task, we show that the networks following the IDIP rule can sustain memory encoding, without disturbing the rest of the network (Figure 2.4F-G). This has previously also been shown with iSTDP (Vogels et al., 2011). However, the key difference is that networks following the IDIP rule exhibit persistent activity of assemblies after encoding (Figure 2.4G-H), unlike iSTDP (Figure 2.22C-D), as well as higher fidelity of recall (Figure 2.4G-H). Persistent activity of the memory cells has been reported in some experiments (Yassin et al. (2010); Ghandour et al. (2019), but see Barron et al. (2017)) and is thought to be involved in memory bindings via Hebbian plasticity, as the increased activity further strengthens the memory trace (Han et al., 2007) and facilitates greater recall (Tonegawa, 2019). Thus, the global network control provided by IDIP may have important functional implications in memory processing and recall.

As highlighted throughout this work, dense connectivity of the inhibitory neurons to their local networks makes them an ideal candidate for network-centered homeostasis. However, more sparsely connected networks would lead the inhibitory neurons to sense a subset of the network and potentially affect a different subset. Thus the final network firing rate would be less constrained (Figure 2.3D, Figure 2.19). This suggests that the network activity is differentially modulated by the IDIP rule depending on the network architecture. In the limit of very sparse networks, IDIP may not even act as a homeostatic mechanism and winner-take-all dynamics may dominate (Figure 2.24). This may be relevant in the prevention of redundancy or in contrast enhancement (Hartline and Ratliff, 1957). Furthermore, inhibitory connectivity is known to vary within (Meyer et al., 2011) and between (Tamamaki and Tomioka, 2010) brain areas. Thus, IDIP rule would have different functional consequences in different brain regions.

The type of inhibitory plasticity we propose here could be implemented in several different ways. Synaptic input integration over the timescale used in our model could be mediated through N-methyl-D-aspartate receptors (NMDARs), which are known to enhance the neuronal computational repertoire (Gasparini and Magee, 2006; Losonczy and Magee, 2006; Poirazi and Mel, 2001; Stuart and Spruston, 2015; Poirazi and Papoutsi, 2020). In the hippocampus, NMDAR were shown to be preferentially localized at feedback synaptic inputs to PV+ cells (Le Roux et al., 2013), and NMDAR knockouts in the PV+ cells were shown to disrupt spatial representation (Korotkova et al., 2010). Moreover, NMDAR have recently been shown to facilitate supralinear dendritic integration from clustered synapses in the PV+ cells (Cornford et al., 2019), a model of which also shows winner-take-all dynamics between competing cell assemblies (Cornford et al., 2019). The inhibitory synaptic plasticity modelled here could occur through a mechanism such as analogue-digital facilitation (ADF) (Alle and Geiger, 2006), where the spike-evoked transmission is graded by presynaptic voltage fluctuations. Indeed, ADF has been reported in inhibitory neurons (Christie et al., 2011; Bouhours et al., 2011; Rowan and Christie, 2017), where it acts through basal Ca^{2+} accumulation and leads to an increase in synaptic release (Christie et al., 2011; Bouhours et al., 2011), as well as Kv channel inactivation (Rowan and Christie, 2017). Alternatively, IDIP could also act through the plasticity of inhibitory neuron intrinsic excitability. In Figure 2.23, we show that our results are qualitatively robust to variations of our learning rule, as long as the plasticity is a function of the synaptic input, or its derivative, to the inhibitory neuron. Thus, in this work we show the functional outcome of a class of models of input-dependent inhibitory plasticity.

Furthermore, our model makes the following experimental predictions. First, we predict that during the optogenetic silencing of the original place map in the Trouche et al. (2016) experiment, the level of inhibition decreases through inhibitory plasticity, and thus the alternative place map emerges. Conversely, we predict that if inhibitory plasticity is blocked during the optogenetic silencing, the alternative place map would still emerge but its activity would be much lower than that of the original place map, as shown in Figure 2.2G-H. Second, we predict that the inhibitory plasticity occurs due to the decrease in the excitatory input to the inhibitory neurons. The role of the excitatory input to the inhibitory neuron plasticity could be experi-

mentally verified by optogenetically stimulating its presynaptic excitatory cells in which case we predict to see a progressive slow increase in the inhibition to the excitatory cells, and vice versa. This could be observed as changes either in the IPSCs magnitude or inhibitory mean spike threshold, depending on whether it is due to synaptic or intrinsic plasticity. Third, we predict that homeostatic recovery should not disrupt pre-existing network structure, as shown in Figure 2.4E. Interestingly, the conservation of network correlations as the network undergoes homeostatic changes was recently reported in a re-analysis of the experimental data from Hengen et al. (2016) by Wu et al. (2020), but this was attributed to synaptic scaling rather than inhibitory plasticity.

Finally, other distinct solutions have been proposed for network-centered homeostasis, such as diffusive neuromodulation via nitric oxide (Sweeney et al., 2015) and GABA spillover (Naumann and Sprekeler, 2020). It is very likely that they all coexist, alongside other, neuron-centered homeostatic mechanisms. Each homeostatic mechanism may control distinct aspects of neural function, as shown in a recent computational study (Wu et al., 2019). Hence, it would be of future interest to study the IDIP rule in conjunction with other homeostatic plasticity rules, such as synaptic scaling (Turrigiano et al., 1998; Desai et al., 2002; Turrigiano and Nelson, 2004; Goel and Lee, 2007; Glazewski et al., 2017), intrinsic plasticity (Desai et al., 1999; Gainey et al., 2018; Lambo and Turrigiano, 2013; Maffei and Turrigiano, 2008), metaplasticity (Bienenstock et al., 1982; Kirkwood et al., 1996; Zenke et al., 2013; Frank et al., 2006), diffusive neuromodulation (Sweeney et al., 2015; Steinert et al., 2008, 2011; Naumann and Sprekeler, 2020), structural plasticity (Yin and Yuan, 2015; Gallinaro and Rotter, 2018) and inhibitory plasticity (Woodin et al., 2003; Maffei et al., 2004, 2006; Keck et al., 2011; Chen et al., 2011; Vogels et al., 2011; van Versendaal et al., 2012; D’amour and Froemke, 2015; Udakis et al., 2019; Clopath et al., 2016; Hennequin et al., 2017; Das et al., 2011; Haas et al., 2006), as their interaction may increase the compensatory repertoire of the networks and endow them with non-trivial emergent properties.

2.5 Methods

2.5.1 Neuron Model

We use the single compartment leaky integrate-and-fire neuron model in our simulations. The model is defined by a resting membrane potential V_{REST} and membrane time constant τ_m . If the membrane potential surpasses the set threshold potential θ_m , it fires a spike and its membrane potential is reset back to V_{REST} . It then enters the refractory period t_{ref} , during which it cannot be stimulated.

The sub-threshold membrane voltage V_i of neuron i follows:

$$\tau_m \frac{dV_i}{dt} = (V_{REST} - V_i) + R(g_{ij}^E(V^E - V_i) + g_{ij}^I(V^I - V_i) + I_{ex}) \quad (2.1)$$

where R is the membrane resistance, $g_{ij}^{E/I}$ are the excitatory and inhibitory synaptic conductances to neuron i from neuron j , $V^{E/I}$ are excitatory and inhibitory reversal potentials and I_{ex} is any other externally supplied current. If neuron i receives input from neuron j , the corresponding synaptic conductance g_{ij} is as follows:

$$\tau_{E/I} \frac{dg_{ij}}{dt} = -g_{ij} + \bar{g}W_{ij}S_j(t) \quad (2.2)$$

where $\tau_{E/I}$ is the synaptic time constant, W_{ij} is the synaptic weight, $\bar{g}$ is a basic unit of synaptic conductance and $S_j(t)$ is the presynaptic spike train.

2.5.2 Input-dependent inhibitory plasticity (IDIP)

The inhibitory plasticity depends on the synaptic input y_i received by each inhibitory neuron i over time.

$$\tau_{IDIP} \frac{dy_i}{dt} = -y_i + \sum_j g_{ij}^E(t) \quad (2.3)$$

where τ_{IDIP} is the input current time constant and g_{ij}^E is the excitatory synaptic conductance of the synapse from an excitatory neuron j to an inhibitory neuron i . Hence the second term is the sum of all of the excitatory input currents to the inhibitory neuron i at time t . When the inhibitory neuron i spikes, all the inhibitory synapses projecting from the presynaptic inhibitory neuron i to the postsynaptic excitatory neuron j are adjusted with respect to y_i as:

$$\Delta w_{ji} = \eta_{IDIP} (y_i(t) - \theta_{in}) S_i \quad (2.4)$$

where η_{IDIP} is the inhibitory learning rate, θ_{in} is the constant target input for each inhibitory neuron and S_i is the inhibitory neuron i spike train. In our simulations, θ_{in} is the same for all inhibitory neurons.

2.5.3 Hippocampal Network Model

We simulate a CA1 network consisting of N_E excitatory neurons and N_I inhibitory neurons. There is full and bidirectional connectivity between the excitatory and inhibitory neurons. The initial synaptic weights from the inhibitory to the excitatory neurons (W_{IE}) are 200 times weaker than the synaptic weights from the excitatory to the inhibitory neurons (W_{EI}). The initial values of the inhibitory synaptic weights are set to ten times weaker than the excitatory weights. The CA1 excitatory neurons are divided in N_g equally sized groups.

Each excitatory neuron in the CA1 network is connected to N_{CA3} CA3 neurons. The excitatory synaptic weights from the CA3 neurons to the CA1 excitatory neurons are made plastic with classical Hebbian plasticity and the inhibitory synaptic weights from the CA1 inhibitory neurons are made plastic with IDIP.

Position-modulated inputs

We model a mouse traversing a 1D annular track with N_{pc} equally spaced place fields. We assume the mouse moves with a constant speed and takes 3 seconds to move from a centre of one place field to a centre of the subsequent place field.

Each CA3 neuron receives input from one unique place field in the form of an external current I_{ex} . The tuning curve T with place field centered at p_0 is defined as:

$$T(p) = \exp\left(-\frac{(p - p_0)^2}{2\sigma_{pc}^2}\right) \quad (2.5)$$

where p is the animal's position and σ_{pc} is the tuning width.

The place field current I_{ex} supplied to a CA3 neuron is then:

$$I_{ex}(p) = A_c T(p) \quad (2.6)$$

where A_c is the current amplitude.

Each CA1 group is tuned to a single unique CA3 neuron, with the same tuning curve T as in Eq. 5. Within group, the tuning amplitude for each neuron is varied. To this end, we sample a vector of $\frac{N^E}{N_g}$ random number from a normal distribution with $\mu = 0$ and $\sigma = 0.05$. We then shuffle this vector for each group and add a single value to the synaptic tuning curve of each neuron in the group.

The excitatory and inhibitory CA1 neurons also receive uniform external currents, I_{ex}^E and I_{ex}^I respectively.

Excitatory plasticity

The excitatory synaptic weights between CA3 and CA1 excitatory neurons follow classical Hebbian plasticity implemented using a symmetrical spike-timing dependent learning rule. A synaptic trace x_i assigned to each neuron i and follows:

$$\tau_{hebb} \frac{dx_i}{dt} = -x_i + \delta(t - t_{spike}) \quad (2.7)$$

where τ_{hebb} is the time constant of the learning window and t_{spike} is the spike time. We also include a homeostatic term which takes into account the sum of all synaptic weights onto the postsynaptic neuron. The synaptic weight w_{ji} from the presynaptic neuron j to the postsynaptic neuron i is updated following:

$$\frac{dw_{ji}}{dt} = (W_{max} - w_{ji})\eta_{hebb}x_ix_j - \eta_{homeo} \left(\sum_i w_{ji} - \theta_{homeo} \right) \quad (2.8)$$

where W_{max} is the maximum excitatory synaptic weight, η_{hebb} is the learning rate of the Hebbian term, η_{homeo} is the learning rate of the homeostatic term and θ_{homeo} is the homeostatic target.

2.5.4 Recurrent Network Model

We simulate a recurrent network consisting of N_E excitatory and N_I inhibitory neurons. The excitatory neurons are randomly connected with a probability $p_{EE} = 0.1$, in line with experiments (Harris and Shepherd, 2015). The inhibitory neurons are randomly connected to the excitatory neurons with a probability $p_{IE} = p_{EI} = 0.2$. Due to the small size of the network, the *in* degree for each excitatory and inhibitory neuron is kept the same. Thus each excitatory neuron has $N_E \times p_{EE}$ excitatory inputs and $4 N_I \times p_{IE}$ inhibitory recurrent synaptic inputs. Furthermore, each inhibitory neuron receives $N_E \times p_{EI}$ excitatory inputs. Due to the small

network size, inhibitory-inhibitory neuron connections are omitted ($p_{II} = 0$). The values for the recurrent weights W_{rec} are sampled from a log normal distribution, in line with experiments (Song et al., 2005; Loewenstein et al., 2011), with $\mu = 1.0$ and $\sigma_{wrec} = 0.1$.

The initial values of the inhibitory synaptic weights are ten times weaker than the excitatory weights ($W_I/W_E = 0.1$). Furthermore, the inhibitory synaptic weight update Δw_{ji} has a multiplicative weight dependence, such that:

$$\begin{aligned} w_{ji} &\leftarrow (W_{max} - w_{ji}) \times \Delta w_{ji} && \text{when } \Delta w_{ji} > 0 \\ w_{ji} &\leftarrow w_{ji} \times \Delta w_{ji} && \text{when } \Delta w_{ji} < 0 \end{aligned} \tag{2.9}$$

where W_{max} is the maximum inhibitory synaptic weight. The inhibitory learning is turned off during the first 15 seconds of the simulation.

Random neuron input generation

Each neuron in the network is randomly connected with a probability of $p_{ex} = 0.2$ to a subset of the N_{ex} external inputs, which are modelled as Poisson process with a mean firing rate F . The values of the input synaptic weights from the external inputs to the excitatory and inhibitory neurons W_{in} are sampled from the same log normal distribution as W_{rec} , with $\mu = 1.0$ and $\sigma_{win} = 0.1$. To generate sufficient network activity, we increase the W_{in} by a factor of 2.5. For the simulation in Figure 2.4A, the values of W_{in} are increased by 50%.

Firing rate rank computation

The same network as above is used in the simulation. We split the activity of the network into 15 second time bins. The first time bin is the network activity before any inhibitory plasticity.

We then compute the Spearman's rank correlation coefficient of the mean excitatory firing rates in the first time bin and each subsequent time bin.

Memory protocol

The same network as above is used in the simulation. After the network activity stabilises (600 s), the recurrent synapses between a selected subset of N^E (12/80) are increased by a factor of 5. After the network reaches steady state again (1200 s), the values of the synaptic weights from the N_{ex}^E to two neurons in the assembly are increased by 50%.

Inhibitory spike-timing dependent plasticity (iSTDP)

With iSTDP (Vogels et al., 2011), the synaptic weight w_{ji} from presynaptic inhibitory neuron i and the postsynaptic excitatory neuron j follows:

$$\begin{aligned} w_{ji} &\leftarrow w_{ji} + \eta_{iSTDP}(x_j - \alpha) && \text{when } \mathbf{presynaptic} \text{ neuron } i \text{ spikes} \\ w_{ji} &\leftarrow w_{ji} + \eta_{iSTDP}x_i && \text{when } \mathbf{postsynaptic} \text{ neuron } j \text{ spikes} \end{aligned} \quad (2.10)$$

where η_{iSTDP} is the learning rate, α is the depression factor and $x_{i/j}$ is the neuron synaptic trace, defined as:

$$\tau_{iSTDP} \frac{dx_{i/j}}{dt} = -x_{i/j} + \delta(t - t_{spike}) \quad (2.11)$$

where τ_{iSTDP} is the learning window time constant and t_{spike} is the spike time.

Code

In all the simulations, we use Euler integration with step size Δt of 1 ms. The Python code for this work is freely available [here](#).

2.6 Simulations parameters summary

Neuron Model		
V_{REST}	- 60 mV	Resting membrane potential
Θ_m	- 50 mV	Membrane potential spiking threshold
R	100 $M\Omega$	Membrane resistance
t_{ref}	2 ms	Absolute refractory period
τ_m	20 ms	Membrane time constant
I_{ex}	0 A	Default external current

Table 2.1: Neuron parameters.

Synapse Model		
$\bar{g}$	1 nS	Basic weight unit
V^E	0 mV	Excitatory reversal potential
V^I	- 80 mV	Inhibitory reversal potential
τ_E	5 ms	Decay constant of AMPA-type conductance
τ_I	10 ms	Decay constant of GABA-type conductance

Table 2.2: Synapse parameters.

IDIP		
τ_{IDIP}	160 ms	IDIP decay constant

Table 2.3: Plasticity parameters.

Simulation		
Δt	1 ms	Integration time step size

Table 2.4: Simulation parameters.

Hippocampal Network Model			
CA3	N_{CA3}	10	Size of the CA3 excitatory population
CA1	N_E	100	Size of the CA1 excitatory population
	N_I	20	Size of the CA1 inhibitory population
	N_g	10	Number of excitatory neuron groups
Position modulated input	N_{pc}	10	Number of place fields
	σ_{pc}	6.54 a.u.	Tuning width
	A_c	200 pA	Current amplitude
External current	I_{ex}^E	0.01 pA	Current to the excitatory neurons
	I_{ex}^I	0.5 pA	Current to the inhibitory neurons
Synaptic weights	W_{IE}	2.0	E to I synaptic weight
	W_{EI}	1e-3	Initial I to E synaptic weight
IDIP	η_{IDIP}	1e-5 $ms^{-1}nS^{-1}$	IDIP learning rate
	θ_{in}	200 nS	Inhibitory neuron target input
Excitatory plasticity	τ_{hebb}	20 ms	Learning window time constant
	W_{max}	1.5	Maximum E to E synaptic weight
	η_{hebb}	1e-3 ms^{-1}	Excitatory Hebbian term learning rate
	η_{homeo}	1e-4 ms^{-1}	Excitatory homeostatic term learning rate
	θ_{homeo}	5.2	Homeostatic target sum of weights

Table 2.5: Hippocampal neural network parameters.

Recurrent Network Model			
Neurons	N_{ex}	100	Size of the input population
	N_E	80	Size of the excitatory population
	N_I	20	Size of the inhibitory population
Connection probability	p_{ex}	0.2	External input to the recurrent
	p_{EE}	0.1	Excitatory to excitatory connectivity
	p_{EI}	0.25	Excitatory to inhibitory connectivity
	p_{IE}	0.25	Inhibitory to excitatory connectivity
	p_{II}	0	Inhibitory to inhibitory connectivity
Ex. input	F	10 Hz	Mean firing rate of external Poisson input
Synaptic weights	μ	1.0	Mean of the sampled distribution
	σ_{win}	0.05	s.d. of the W_{in} sampled distribution
	σ_{wrec}	0.05	s.d. of the W_{rec} sampled distribution
	W_I/W_E	0.1	Initial inhibitory/ excitatory weight ratio
IDIP	η_{IDIP}	$1e-7 \text{ ms}^{-1} nS^{-1}$	IDIP learning rate
	θ_{in}	550 nS	Inhibitory neuron target input
	W_{max}	1.0	Maximum I to E synaptic weight
iSTDP	τ_{iSTDP}	20 ms	Learning window time constant
	η_{iSTDP}	$5e-5 \text{ ms}^{-1}$	iSTDP learning rate
	α	0.2	Depression factor

Table 2.6: Recurrent neural network parameters.

2.7 Supplementary figures

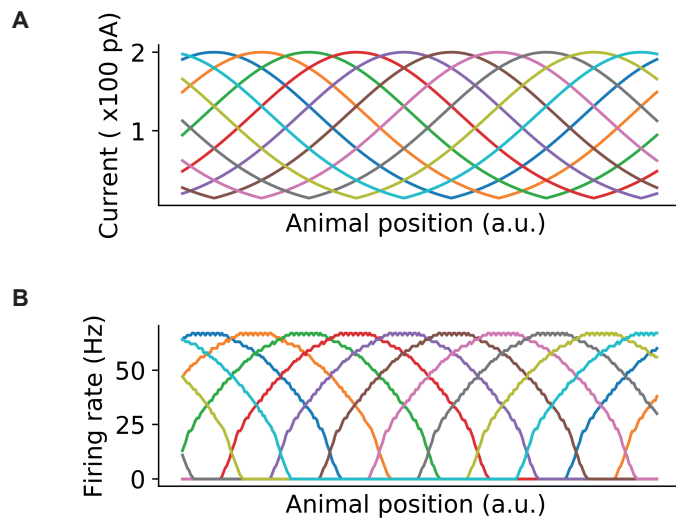


Figure 2.5: **A** The external current supplied to the CA3 neurons. Different colours indicate different place fields. **B** The CA3 neuron firing rates. Different colours indicate different CA3 neurons tuned to different place fields, as in **A**.

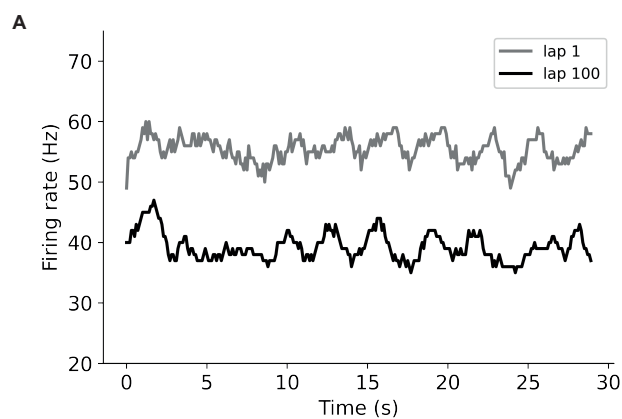


Figure 2.6: **A** Firing rate of the CA1 inhibitory neurons in the first (grey) and last (black) lap of the simulation.

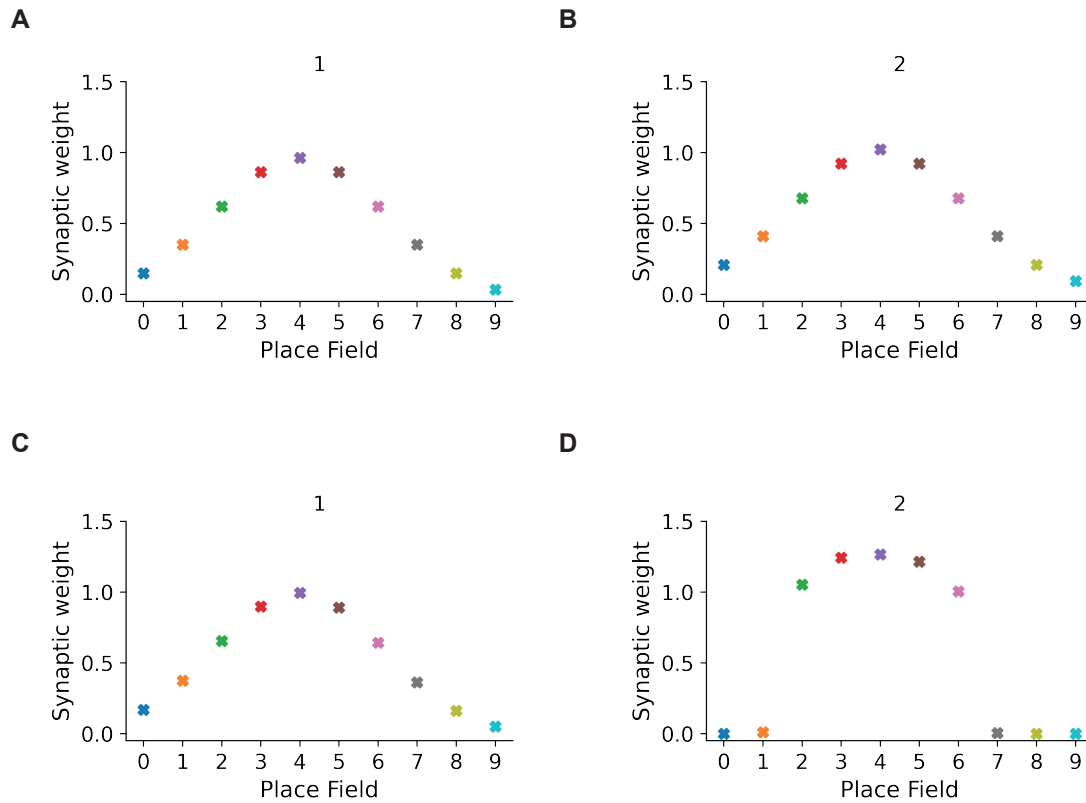


Figure 2.7: **A** The initial CA3 to CA1 synaptic weights of neuron 1, as in Figure 2.1C. Different colours indicate synapses from CA3 neurons with different place fields, as in S1. **B** The initial CA3 to CA1 synaptic weights of neuron 2, as in Figure 2.1C. **C** The final CA3 to CA1 synaptic weights of neuron 1, as in Figure 2.1F. **D** The final CA3 to CA1 synaptic weights of neuron 2, as in Figure 2.1F.

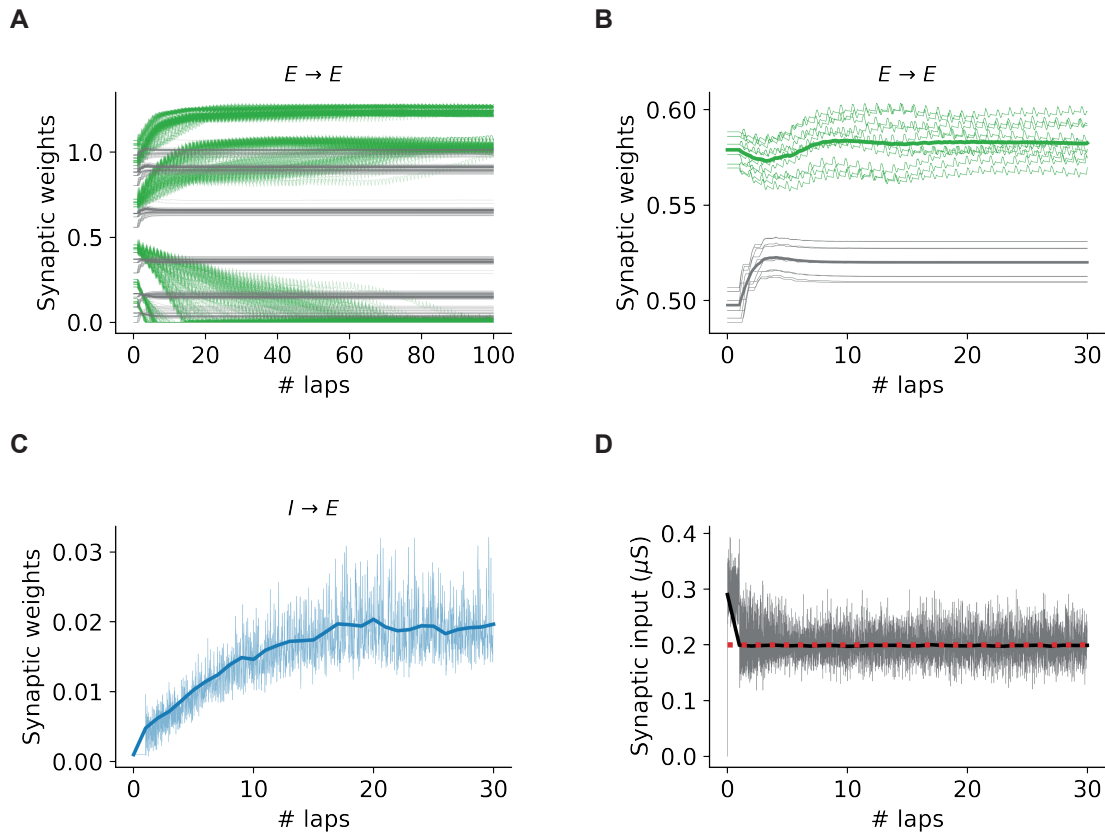


Figure 2.8: **A** CA3 to CA1 synaptic weights of an example place (green) and silent (grey) cell as a function of lap number. **B** Mean CA3 to CA1 synaptic weights of an all place (green) and silent (grey) cell as a function of lap number. The bold lines denote the mean over all active cells. **C** CA1 inhibitory synaptic weights as a function of time. The bold lines denote the mean value per lap. **D** The synaptic input to the inhibitory neurons (grey line) as a function of time. The bold black line denotes the mean per lap. The red dotted line denotes the target input.

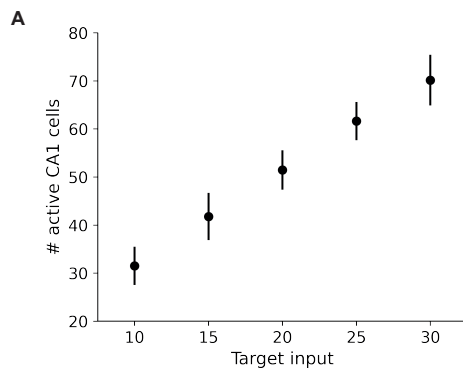


Figure 2.9: **A** The number of active place cells as a function of target input value. The error bars denote the mean over 20 networks.

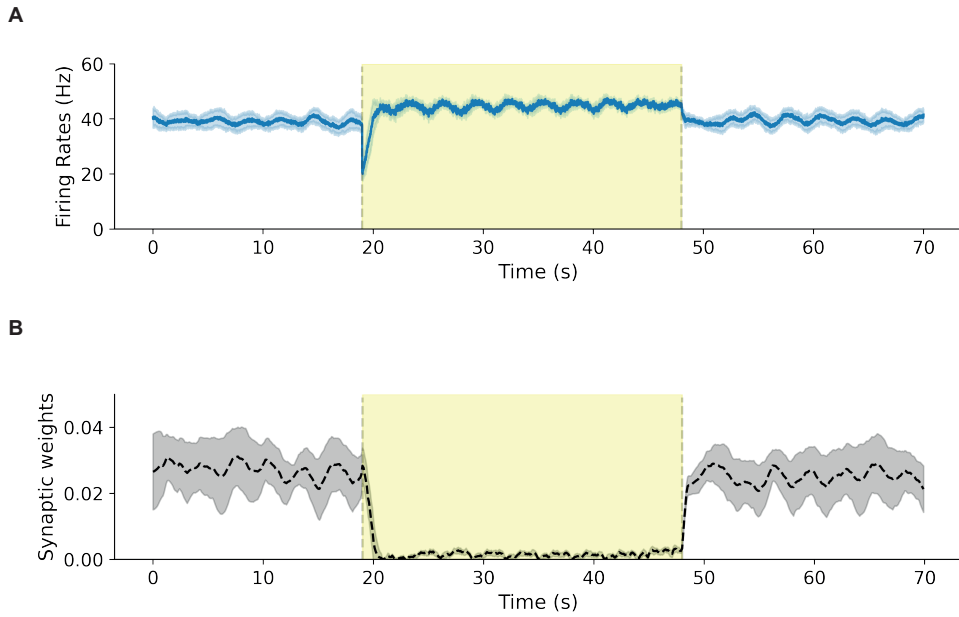


Figure 2.10: **A** Evolution of the inhibitory firing rates during the silencing protocol as described in Figure 2.2A. **B** Evolution of the inhibitory synaptic weights during the silencing protocol as described in Figure 2.2A.

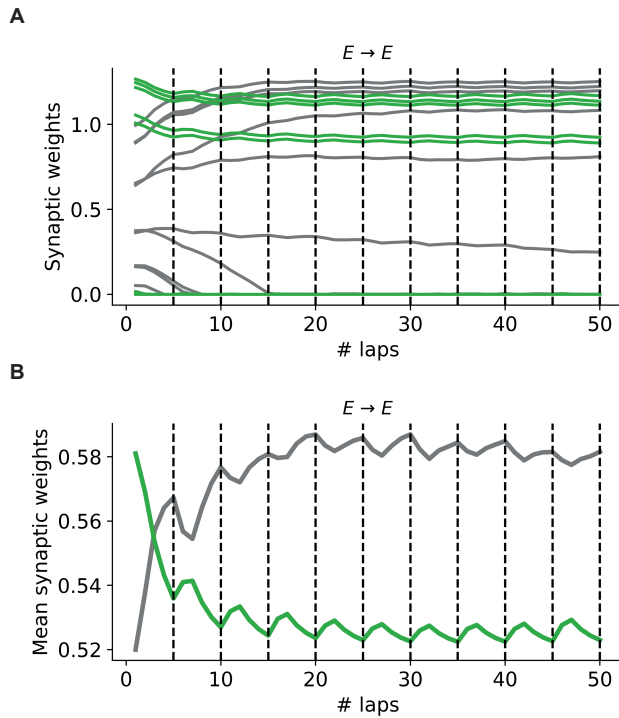


Figure 2.11: **A** The evolution of all the synaptic weights from CA3 to CA1 for an example place (green) and silent (grey) cell during the consolidation protocol as described in Figure 2.2I. Dashed vertical lines mark testing laps. **B** Same as in **A**, but for the mean synaptic weights an example place (green) and silent (grey) cell.

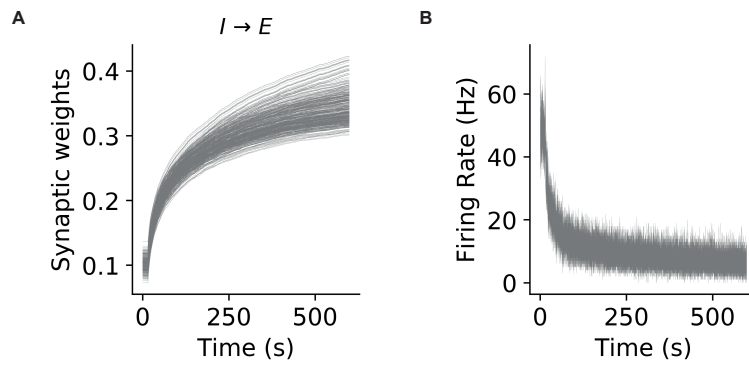


Figure 2.12: Recurrent network. **A** All the inhibitory synaptic weights. **B** All the excitatory firing rates.

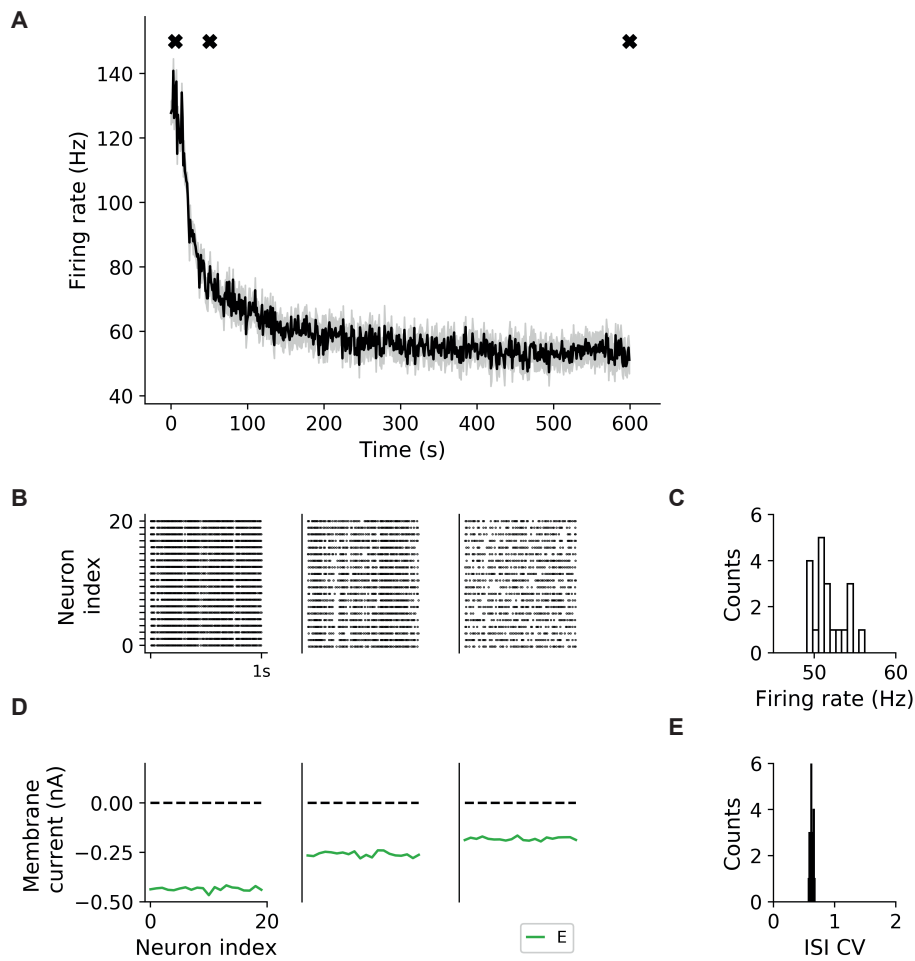


Figure 2.13: **A-E** Same as in Figure 2.3**B,E,F,C** but for the inhibitory neurons only.

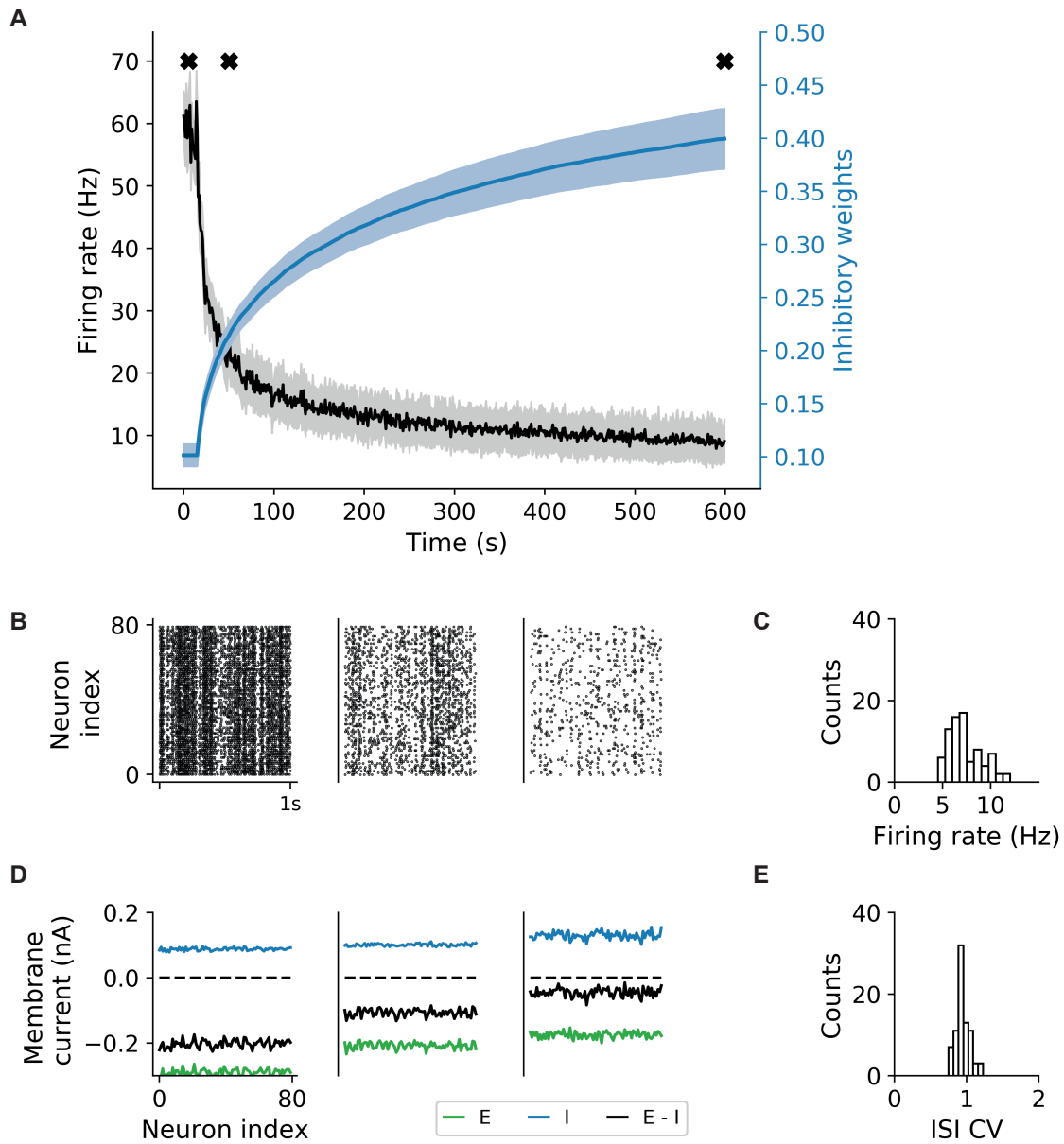


Figure 2.14: **A-D** Same as in Figure 2.3**B,E,F,C** but for the network with fixed inhibitory-inhibitory connections. The inhibitory neurons are connected with a probability $p_{II} = 0.2$, with the in-degree kept constant.

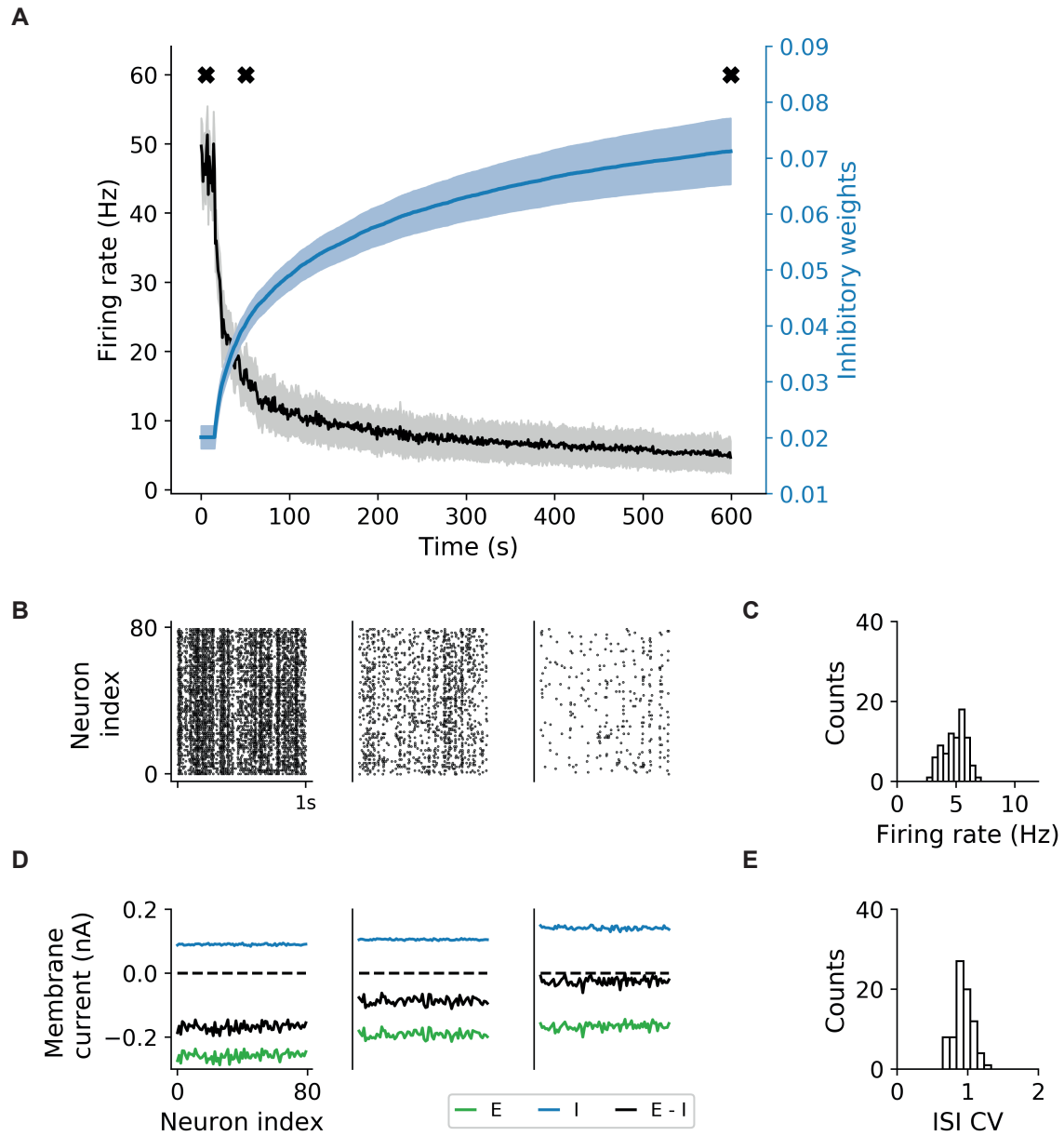


Figure 2.15: **A-D** Same as in Figure 2.3**B,E,F,C** but for a bigger network ($N_E=400$, $N_I=100$). Here we use $1/N$ scaling for W_{rec} .

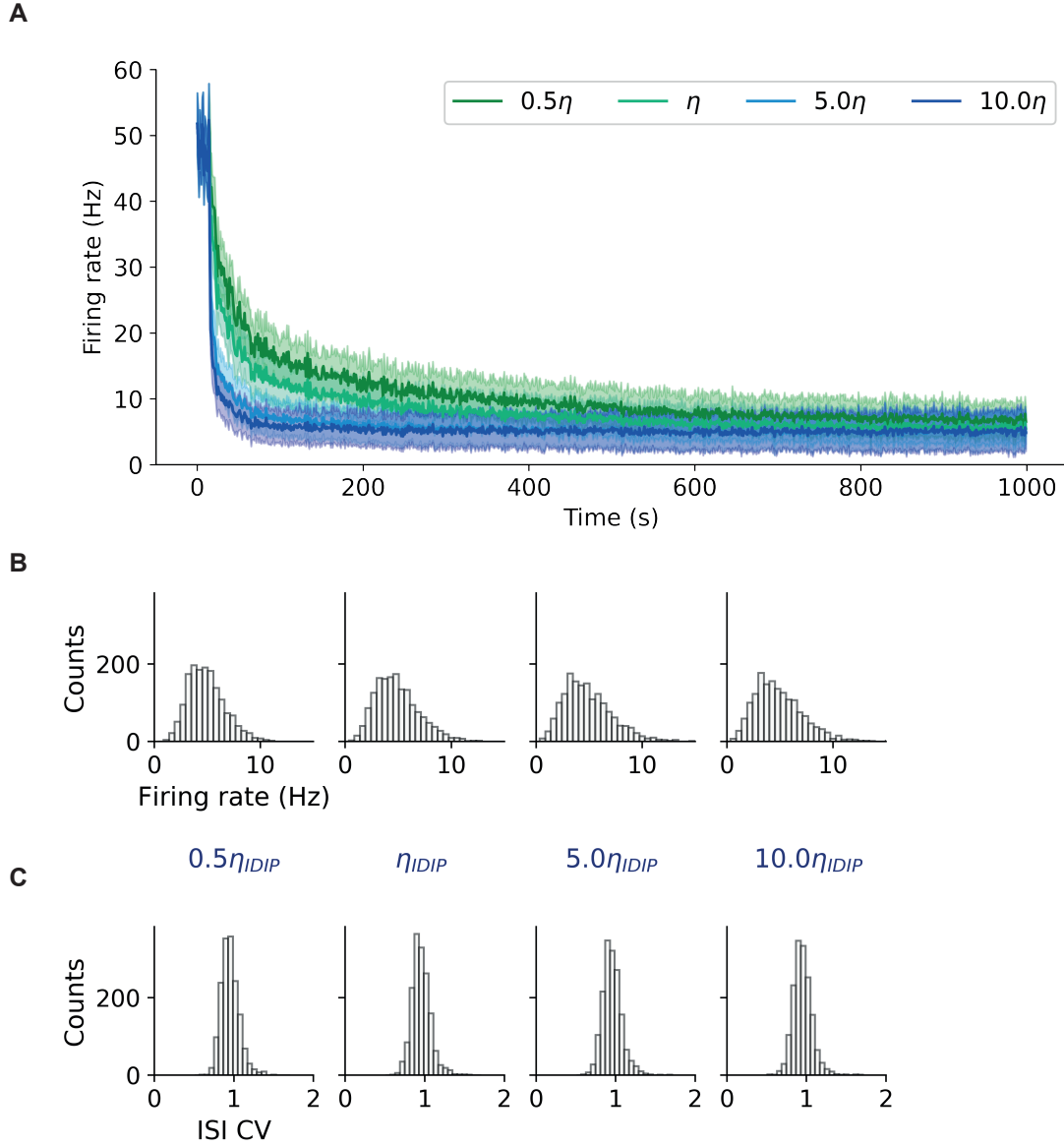


Figure 2.16: **A** Firing rate as a function of time with varying learning rate η_{IDIP} . The shaded area denotes the standard deviation of the firing rates over all the neurons in the network. **B** Firing rate distribution of networks with different learning rates η_{IDIP} after inhibitory learning. Here we include data from 20 different networks. **C** ISI CV of networks with different learning rates after inhibitory learning. Here we include data from 20 different networks.

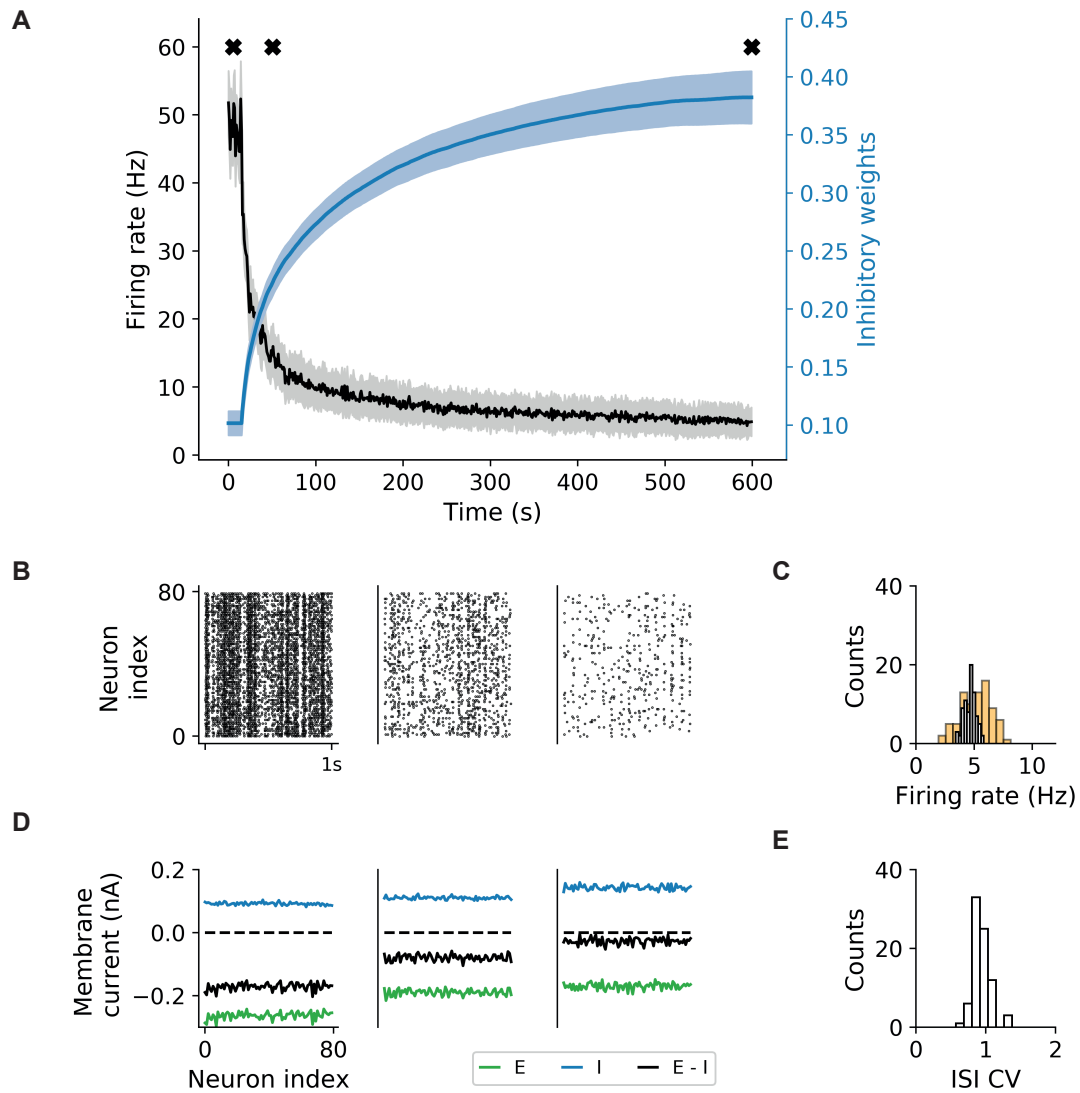


Figure 2.17: **A-D** Same as in Figure 2.3**B,E,F,C** but for the network following iSTDp (Vogels et al., 2011). **C** The pale orange histogram indicates the distribution of firing rates with IDIP in the same network.

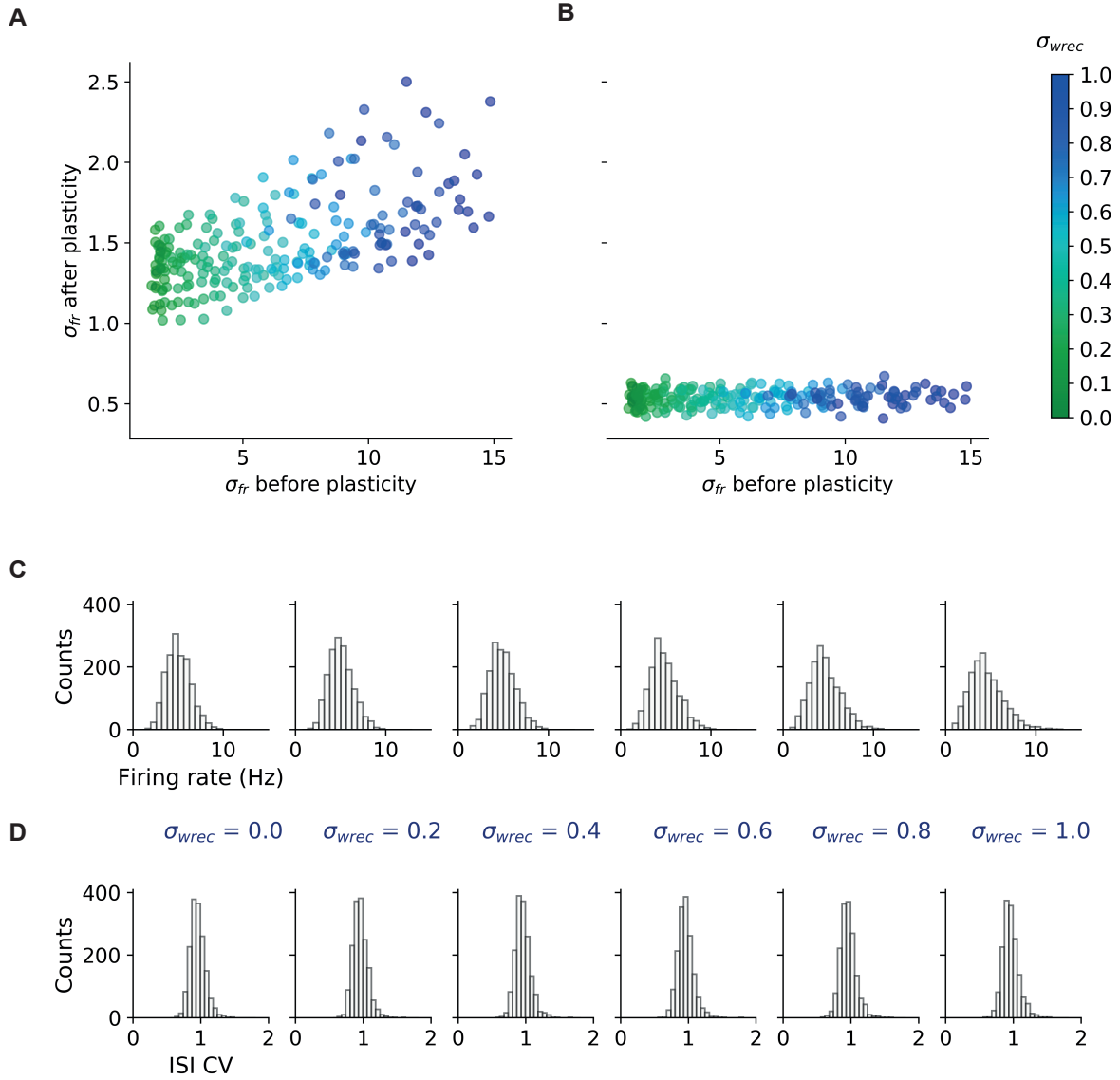


Figure 2.18: **A** Standard deviation of the excitatory firing rates σ_{fr} before (x-axis) and after (y-axis) inhibitory learning for the networks following IDIP (Pearson correlation coefficient $r = 0.59$, p-value $= 2.8e-22$). Different standard deviation of the W_{rec} sampling distribution (σ_{wrec}) values are color-coded. **B** Same as in **A** but for networks following iSTDP (Vogels et al., 2011) (Pearson correlation coefficient $r = 0.08$, p-value $= 0.22$). **C** Final firing rate distribution of networks following IDIP with different W_{rec} sampling distribution standard deviation values σ_{wrec} . Here we include data from 20 different networks. **D** Same as in **C** for ISI CV.

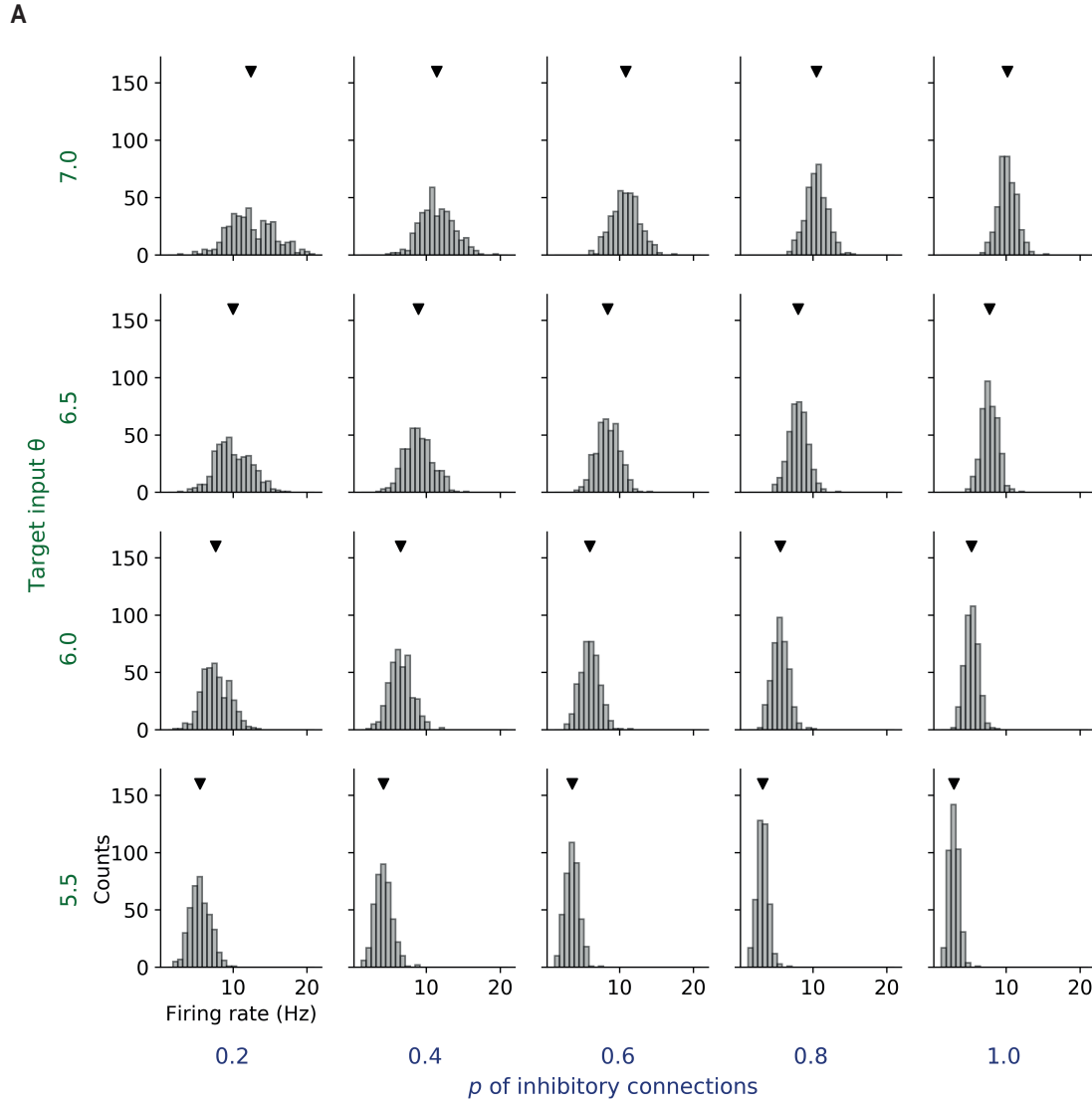


Figure 2.19: **A** The distribution of the firing rates after IDIP as a function of different values of inhibitory probability of connections (horizontal axis) and the target input θ_i (vertical axis). Each histogram contains data from 20 different networks. The means of each distribution are denoted with black triangles. As the target input value increases (y axis), we observe an increase in the mean firing rate (black triangles) and an increase in firing rate diversity, as the network is less constrained. As the inhibitory connectivity increases (x axis), the excitatory population subsets that each inhibitory neurons is connected to increasingly overlap, and thus there is a decrease in the diversity of the excitatory firing rates.

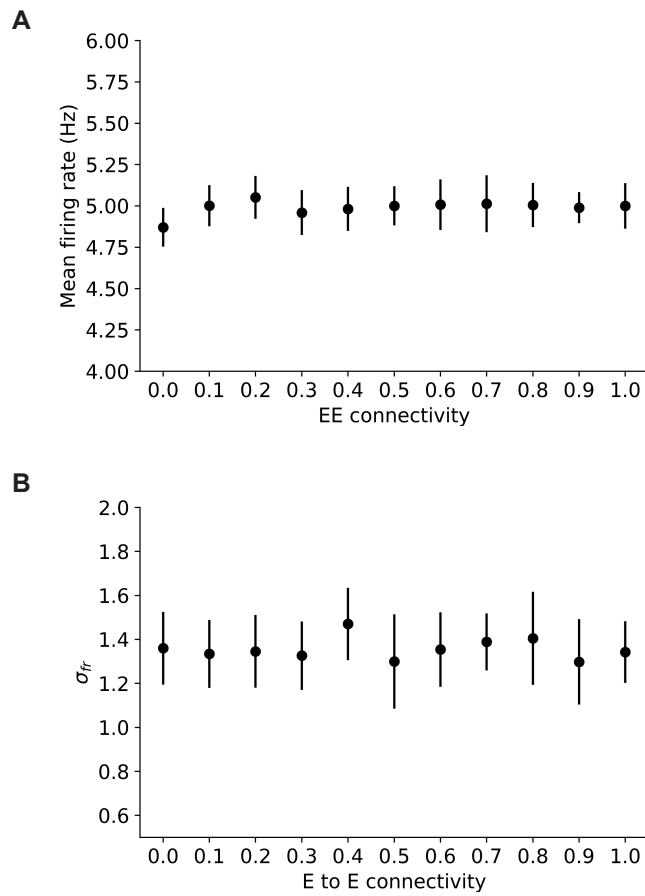


Figure 2.20: **A** The mean firing rate of networks with varying excitatory connectivity. The error bars denote the standard deviation over 20 networks.

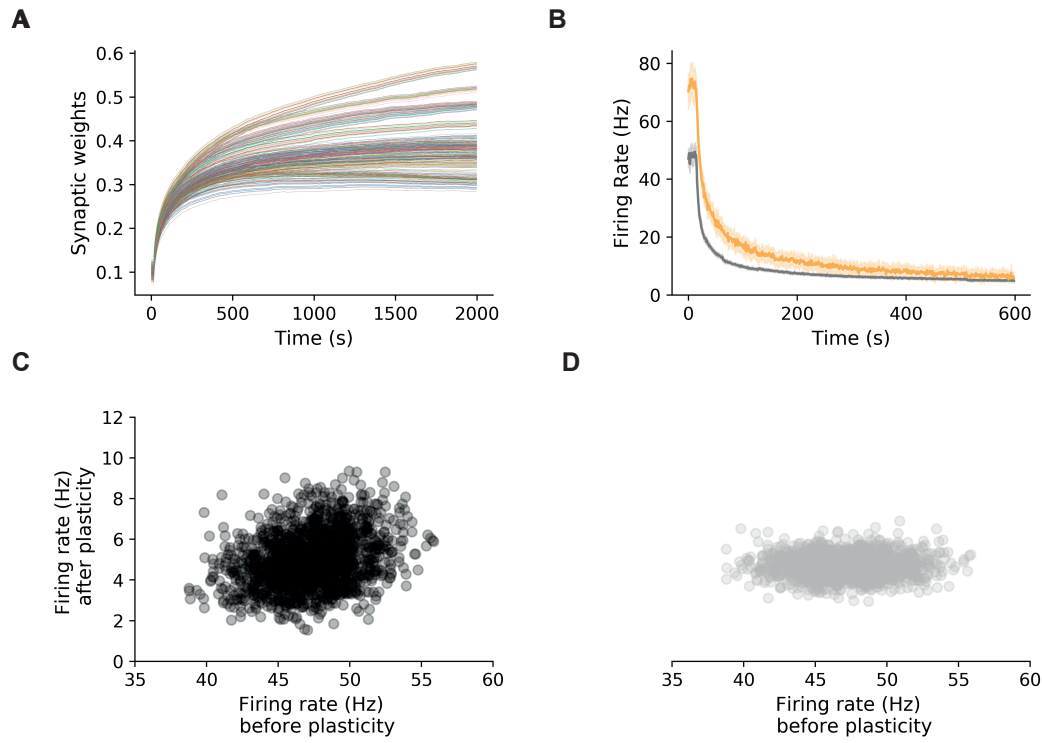


Figure 2.21: **A-B** Recurrent network where a subset of neurons receives increased input, as in Figure 2.4A. . **A** The evolution of all the inhibitory synaptic weights **B** Same as Figure 2.4B but in a network following iSTDP (Vogels et al., 2011). **C-D** Recurrent network as in Figure 2.3A, with no neurons receiving increased external input. **C** The excitatory firing rate before (x-axis) versus after (y-axis) inhibitory learning in the network following IDIP. (Pearson correlation coefficient $r = 0.29$, p-value = $5.09e-32$) **D** The excitatory firing rate before (x-axis) versus after (y-axis) inhibitory learning in the network following iSTDP (Vogels et al., 2011) (Pearson correlation coefficient $r = 0.005$, p-value = 0.84).

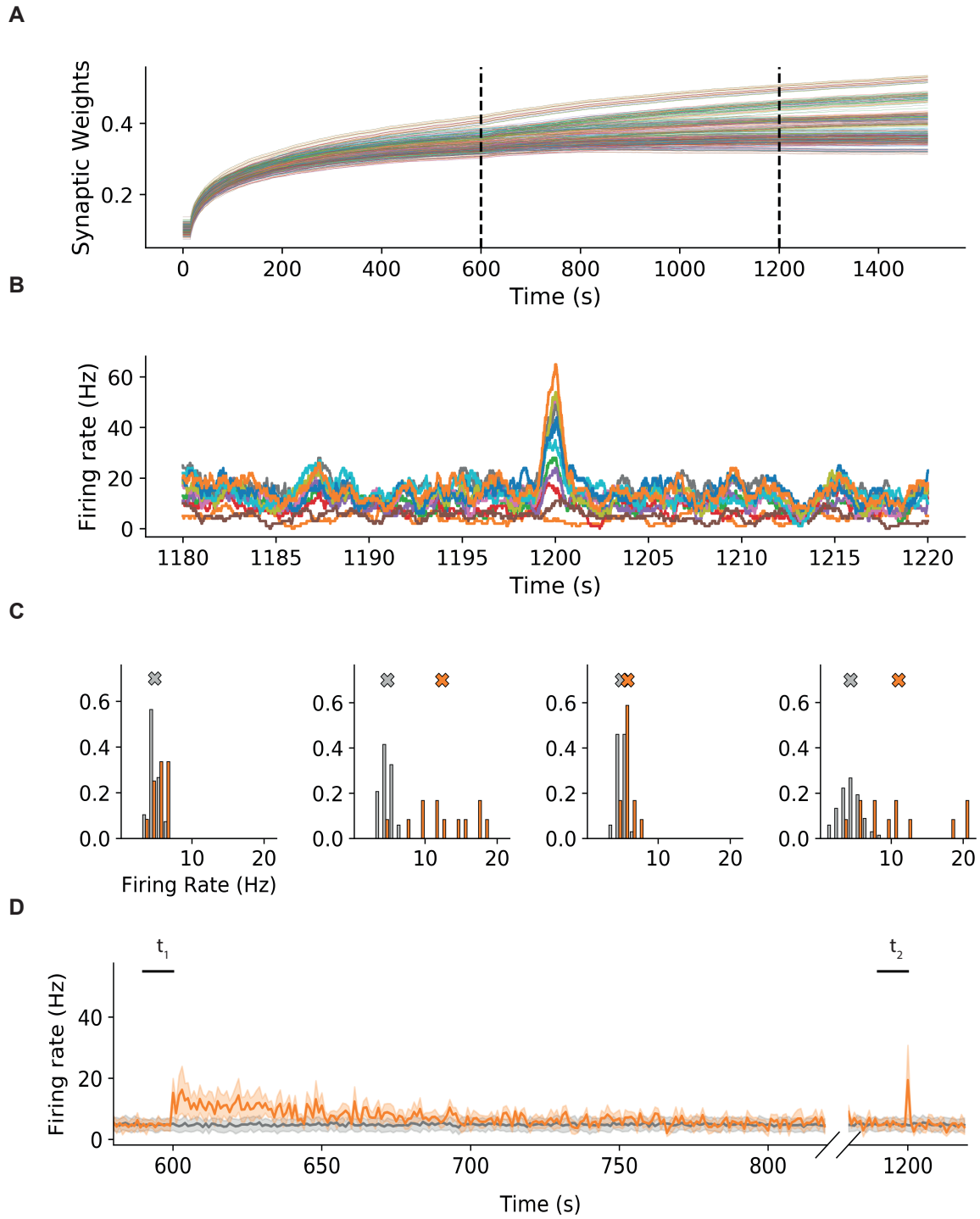


Figure 2.22: **A** All the inhibitory synaptic weights during the associative memory task (Figure 2.4F). The dashed vertical lines indicate memory encoding and recall respectively. **B** The assembly firing rates during recall. **C-D** Same as Figure 2.4G-H but in a network following iSTDP (Vogels et al., 2011).

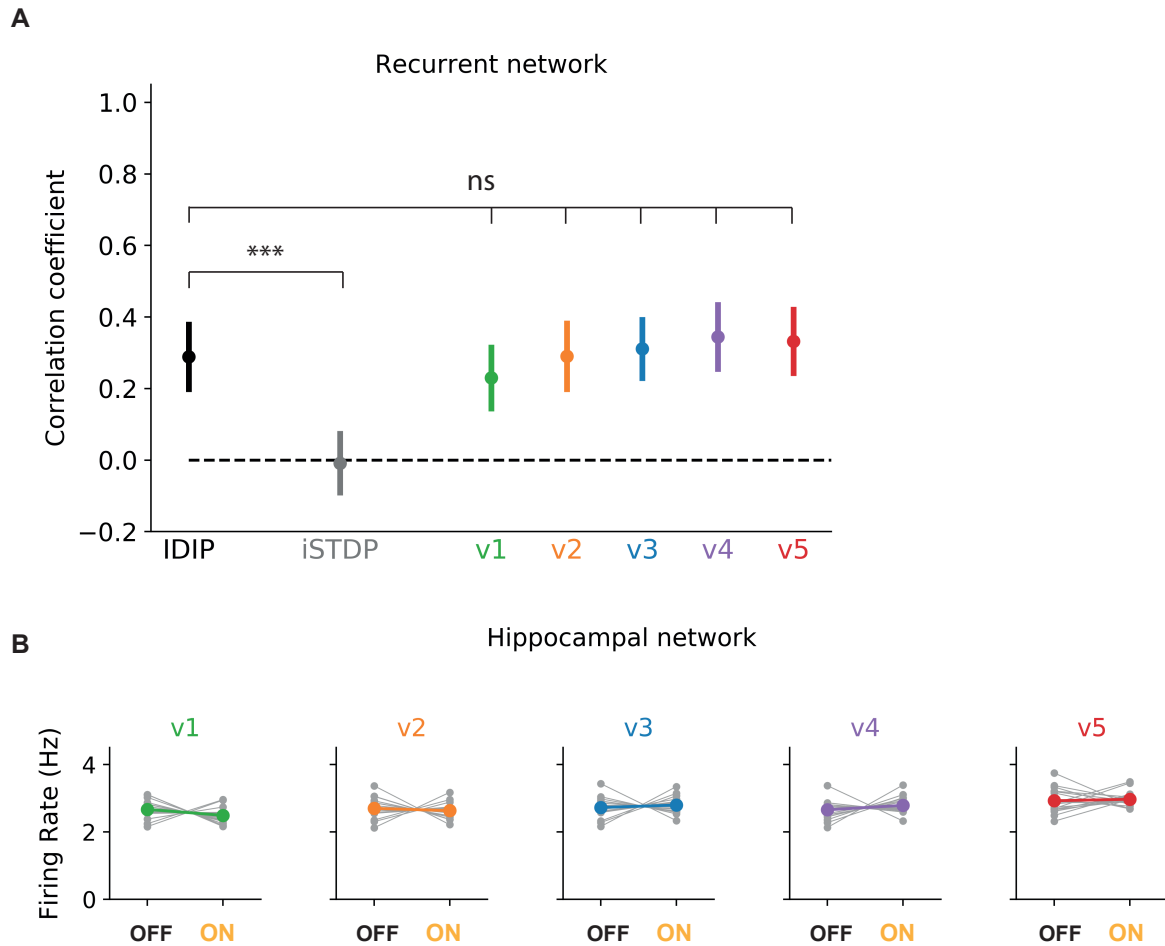


Figure 2.23: **A** The final firing rate rank correlation coefficient, calculated as in Figure 2.4E, for the recurrent networks following different IDIP learning rule variants (outlined in the supplementary methods). The correlation coefficient values of the networks following IDIP and iSTDP are significantly different (student t-test, p value = $7e-12$). The correlation coefficient values of the networks following IDIP and IDIP learning rule variants are not significantly different (student t-test, p values = 0.07 (v1), 0.96 (v2), 0.47 (v3), 0.09 (v4), 0.18 (v5)). **B** The mean firing rate of the place map before ('OFF') and during ('ON') silencing for the hippocampal networks following IDIP learning rule variants. Gray circles indicate individual trials. Colour-coded circles indicate average over all 20 trials.

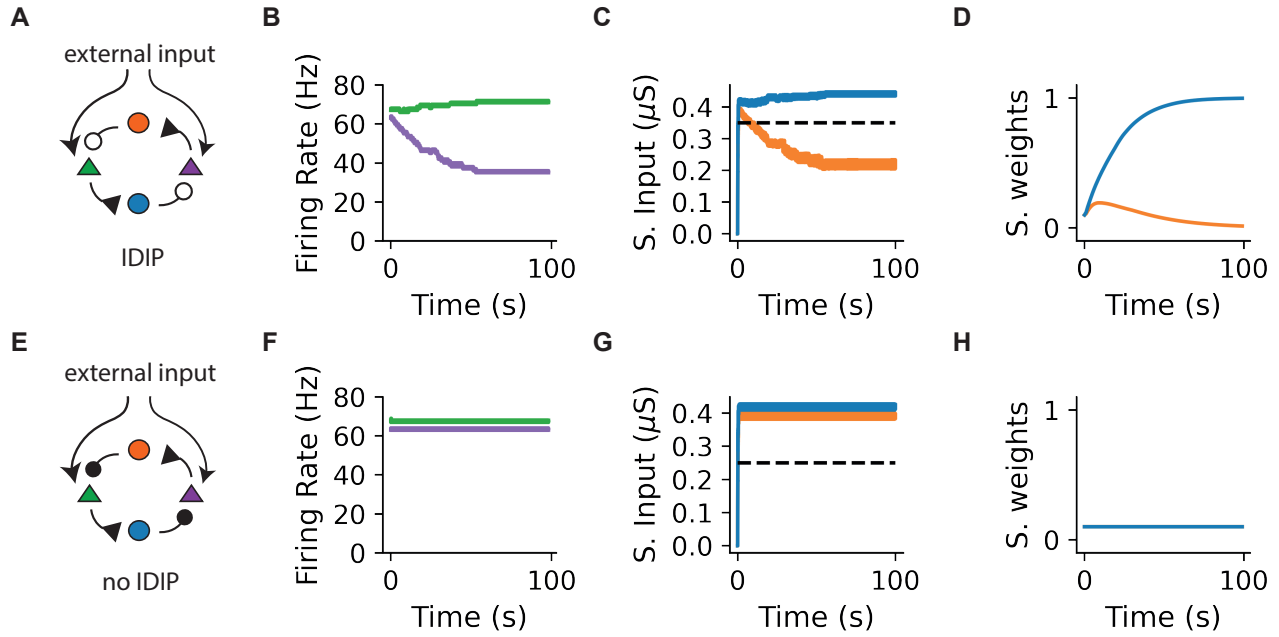


Figure 2.24: **A** Diagram of a network with sparse inhibitory connectivity. Two excitatory (triangles) and two inhibitory (circles) neurons are simulated. One of the excitatory neurons (green) receives a slightly (5%) stronger external input. Thus one inhibitory neuron (blue) receives synaptic input from the excitatory neuron that is more active and projects onto the other excitatory neuron (purple). The other inhibitory neuron (orange) receives synaptic input from the less active excitatory neuron (purple) and projects onto the more active (orange) excitatory neuron. **B** The evolution of neuronal firing rates for the excitatory cells, color coded as in **A**. The initially small activity difference is amplified by the inhibitory plasticity, leading to increased activity of one excitatory neuron (green) and a decrease in the other (purple). **C** The synaptic input to the inhibitory neurons as a function of time, color coded as in **A**. The dashed line denotes the target input. **D** The evolution of the inhibitory (I to E) synaptic weights, color coded as in **A**. **E-H** The same as in **A-D** but without any plasticity.

2.8 Supplementary methods

2.8.1 Various learning rule implementations

To show that our results is robust to the specific implementation of input dependent inhibitory plasticity (IDIP), we devise several learning rule variants (see Figure 2.23).

Original learning rule

In the learning rule implementation used throughout this paper, the inhibitory plasticity depends on the low pass filtered synaptic input y_i received by each inhibitory neuron i .

$$\tau_{IDIP} \frac{dy_i}{dt} = -y_i + \sum_j g_{ij}^E(t) \quad (2.12)$$

$$\Delta w_{ji} = \eta_{IDIP} (y_i(t) - \theta_{in}) S_i(t) \quad (2.13)$$

where τ_{IDIP} is the input current time constant, g_{ij}^E is the excitatory synaptic conductance of the synapse from an excitatory neuron j to an inhibitory neuron i , $S_i(t)$ is the spike train of the inhibitory neuron i , η_{IDIP} is the inhibitory learning rate and θ_{in} is the constant target input for each inhibitory neuron. The values for each parameter can be found in the parameters summary table in the main text. The adjusted parameters for each learning rule variant can be found in the supplementary tables 2.7 and 2.8.

Variant 1

In the first variant ('v1') of the learning rule, we add a voltage term to Eq 12.

$$\frac{dy_i}{dt} = -\frac{y_i}{\tau_{IDIP}} + \sum_j g_{ij}^E(t) S_j(t) (V^E - V_i) \quad (2.14)$$

where V_i is the sub-threshold membrane voltage of the inhibitory neuron i and V^E is the excitatory reversal potential.

Variant 2

In the second variant ('v2') of the learning rule, we replace the spike train $S_i(t)$ term in Eq 13 by a synaptic trace x_i .

$$\tau_{st} \frac{dx_i}{dt} = -x_i + S_i(t) \quad (2.15)$$

$$\Delta w_{ji} = \eta_{IDIP} (y_i(t) - \theta_{in}) x_i(t) \quad (2.16)$$

where τ_{st} is the synaptic trace time constant.

Variant 3

In the third variant ('v3') of the learning rule, we add inhibitory synapses to the inhibitory neurons in our network and include their synaptic input contribution in Eq 12.

$$\tau_{IDIP} \frac{dy_i}{dt} = -y_i + \sum_j g_{ij}^E(t) - \sum_j g_{ij}^I(t) \quad (2.17)$$

where g_{ij}^I are the synaptic conductances of the inhibitory synapses to the inhibitory neuron i .

Variant 4

In the fourth variant ('v4') of the learning rule, we replace the inhibitory synaptic plasticity by plasticity in the inhibitory neuron firing threshold θ_i .

$$\Delta\theta_i = \eta_{IDIP}(y_i(t) - \theta_{in}) \quad (2.18)$$

Variant 5

In the fifth variant ('v5') of the learning rule, we replace the inhibitory neuron target input θ_{in} with the inhibitory neuron target firing rate θ_x . We define a running firing rate estimator x_i^{est} .

$$\frac{dx_i^{est}}{dt} = -\frac{x_i^{est}}{\tau_{est}} + S_i(t) \quad (2.19)$$

where τ_{est} is the integration time constant, as done in Wu et al. (2020). The inhibitory weights are updated as:

$$\Delta w_{ji} = \eta_{IDIP} \left(\frac{x_i^{est}(t)}{\tau_{est}} - \theta_x \right) S_i(t) \quad (2.20)$$

2.8.2 Supplementary parameter tables

Hippocampal network			
v1	η_{IDIP}	$4\text{e-}4 \text{ ms}^{-1} nS^{-1}$	IDIP learning rate
	θ_{in}	-1.0 nA	Inhibitory neuron target input
v2	η_{IDIP}	$5\text{e-}7 \text{ ms}^{-1} nS^{-1}$	IDIP learning rate
	τ_{st}	20 ms	Synaptic trace time constant
v4	η_{IDIP}	$5\text{e-}6 \text{ ms}^{-1} mV^{-1}$	IDIP learning rate
	W_{EI}	$1\text{e-}2$	Initial I to E synaptic weight
v5	η_{IDIP}	$1\text{e-}4 \text{ ms}^{-1}$	IDIP learning rate
	τ_{est}	20 ms	Firing rate estimator time constant
	θ_x	80 Hz	Inhibitory neuron target firing rate

Table 2.7: The parameters used in Figure 2.23B

Recurrent network			
v1	η_{IDIP}	$4\text{e-}6 \text{ ms}^{-1}nS^{-1}$	IDIP learning rate
	θ_{in}	-3.0 nA	Inhibitory neuron target input
v2	η_{IDIP}	$5\text{e-}9 \text{ ms}^{-1}nS^{-1}$	IDIP learning rate
	τ_{st}	20 ms	Synaptic trace time constant
v4	η_{IDIP}	$5\text{e-}10 \text{ ms}^{-1}mV^{-1}$	IDIP learning rate
	W_I/W_E	0.15	Inhibitory/excitatory weight ratio
v5	η_{IDIP}	$1\text{e-}6 \text{ ms}^{-1}$	IDIP learning rate
	τ_{est}	20 ms	Firing rate estimator time constant
	θ_x	100 Hz	Inhibitory neuron target firing rate

Table 2.8: The parameters used in Figure 2.23A.

Recurrent Network Model			
Neurons	N_E	2	Size of the excitatory population
	N_I	2	Size of the inhibitory population
Ex. input	I_E	200/210 pA	External current to the excitatory neurons
	I_I	20 pA	External current to the inhibitory neurons
Synaptic weights	W_{IE}	7.0	E to I synaptic weight
	W_{EI}	0.1	Initial I to E synaptic weight
IDIP	η_{IDIP}	$1.5\text{e-}5 \text{ ms}^{-1}nS^{-1}$	IDIP learning rate
	θ_{in}	350 nS	Inhibitory neuron target input

Table 2.9: The parameters used in Figure 2.24.

Chapter 3

Feedback guides learning in biological circuits

This chapter mechanistically studies the model originally proposed in Feulner et al. (2022), with a specific focus on fine-tuning task perturbation through a combination of plasticity and feedback control. The investigation examines how feedback control of ongoing recurrent network dynamics can approximate the optimal first-order gradient with respect to network activities, allowing for rapid, ongoing movement correction. Additionally, it demonstrates that network adaptation to task perturbation using a local, biologically plausible learning rule becomes both more accurate and efficient when feedback control is integrated during learning. This efficiency arises from the temporal decoupling of network dynamics and the injection of adaptive, second-order gradients. The findings suggest that feedback control may guide credit assignment in biological circuits, enabling both rapid and efficient learning in the brain.

This chapter has been submitted to a conference and is currently in review.

Feedback control guides credit assignment in biological circuits

Klara Kaleb¹, Barbara Feulner¹, Juan A. Gallego¹, Claudia Clopath¹

¹Imperial College London, Department of Bioengineering, London, UK

Abstract

How do brain circuits learn to guide our behaviour? While significant strides have been made in understanding learning in artificial neural networks, applying this knowledge to biological networks remains challenging. For instance, while backpropagation is known to perform accurate credit assignment of error, how a similarly powerful process can be realized within the constraints of biological circuits remains largely unclear. One of the major challenges is that the brain’s extensive recurrent connectivity requires the propagation of error through both space and time, a problem that is notoriously difficult to solve in vanilla recurrent neural networks. Moreover, the extensive feedback connections in the brain are known to influence and control forward network activity, but the interaction between the feedback-driven activity changes with weight-based learning is less well understood. Inspired by previous work on motor control models, this work investigates the mechanistic properties of networks pre-trained with feedback control on a stereotypical motor task. We show that the feedback control of the ongoing recurrent network dynamics approximates the optimal first-order gradient w.r.t. the network activities, allowing for rapid, ongoing movement correction. Moreover, we show that the network adaptation to task perturbation using a local, biologically plausible learning rule is both more accurate and more efficient with feedback control during learning, due to the decoupling of the network dynamics *and* injection of adaptive, second-order gradient into the network dynamics. Thus, our results suggest that feedback control may *guide* credit assignment in biological circuits, enabling both rapid and efficient learning in the brain.

3.1 Introduction

Despite the “unreasonable effectiveness” of the backpropagation (BP) algorithm (Rumelhart et al., 1986), learning in the brain must leverage a different solution (Lillicrap and Santoro, 2019; Lillicrap et al., 2020). For instance, the brain’s extensive recurrent connectivity (Felleman and Van Essen, 1991; Douglas and Martin, 2004) requires propagation of error through both space *and* time. Learning in vanilla recurrent neural networks (RNNs) with Backpropagation through time (BPTT) (Werbos, 1990; Mozer et al., 1995; Robinson and Fallside, 1987) is known to be difficult due to the vanishing and exploding gradient problem (Pascanu et al., 2013; Hochreiter and Schmidhuber, 1997). This challenge is exacerbated under biological constraints, which require learning rules with significant approximations of the full temporal memory (Marschall et al., 2020; Bellec et al., 2020; Murray, 2019; Liu et al., 2021), and often lead to an undesirable generalization gap (Liu et al., 2022).

In addition to intra-layer recurrence, the brain also features extensive feedback connections from top-down sources (Peters and Payne, 1993; Peters et al., 1994) that can modulate network activity in lower layers (Przybylszewski, 1998; Girard et al., 2001; De Pasquale and Sherman, 2011; Briggs and Usrey, 2011; Jordan and Keller, 2020; Keller et al., 2020) and improve bottom-up information processing (Gilbert and Li, 2013; Manita et al., 2015; Fyall et al., 2017). This multiplicity of simultaneously active pathways contrasts vanilla RNNs, where feedback connections are used exclusively for gradient propagation.

It is precisely this simultaneous integration of several streams of information that is thought to mediate ongoing movement correction in both humans and non-human primates (Thoroughman and Shadmehr, 2000; Perich et al., 2018; Krakauer et al., 2000). For example, when subjects receive distorted feedback during a learned reaching movement, such as an unexpected force acting on the limb, they are able to make movement corrections already at the first trial (Thoroughman and Shadmehr, 2000). Moreover, they rapidly adapt to the mismatch between the observed and expected movement trajectory upon further perturbed trials, requiring no more online corrections (Krakauer et al., 2000). Recent data-driven modelling work (Feulner et al., 2022) proposed an elegant solution to this problem by integrating feedback control into the

recurrent network dynamics, and pre-training the whole network on a stereotypical motor task. The resulting networks showed rapid adaptation to task perturbation during ongoing movement and further improvement in the task performance upon persistent perturbation using a *feedback-driven, local learning rule*, and thus reproduce key behavioural and neurophysiological findings. However, the relationship between this local feedback-driven learning and gradient descent, or any of its approximations, remains unclear.

In this work, we investigate the mechanistic properties of such recurrent networks pre-trained with feedback control on a stereotypical motor task. Our key findings are listed as follows:

1. Feedback control allows for approximate inference learning in the activity space 3.3.2.
2. Feedback control enables increased accuracy of approximate, local learning rules in the recurrent layer due to the “decoupling” of the network from its past activity 3.3.4.
3. Feedback control enables more efficient weight updates during task adaptation due to the injection of adaptive, second-order gradient into the network dynamics 3.3.5.

Related work

Most modelling studies, the role of feedback connections is restricted to gradient signal propagation only (Bellec et al., 2020; Murray, 2019; Payeur et al., 2021; Liu et al., 2021). The need for separation seems clear: in the studies where this restriction is lifted, the plasticity mechanisms involved often require tight coordination (Scellier and Bengio, 2017; Whittington and Bogacz, 2017; Sacramento et al., 2018; Payeur et al., 2021; Podlaski and Machens, 2020) that may be hard to implement and maintain in the brain. Furthermore, the relative influence of feedback may still be kept small to avoid interference with the forward inference.

Nevertheless, several other computational studies have shown that a stronger influence of feedback on the network activity may actually aid network training (Gilra and Gerstner, 2017; Denève et al., 2017; Alemi et al., 2018; Bourdoukan and Deneve, 2015; Meulemans et al., 2021,

2022a,b). We highlight (Meulemans et al., 2022b), where the authors took this approach a step further by formally deriving the relationship between gradient-based learning and optimal control through feedback activity influence for equilibrium systems, arguing that this alleviates the temporal storage load of the intermediate states required for BPTT-based learning. However, despite the success on a number of difficult problems, it remains unknown how well these findings extend to out-of-equilibrium dynamics commonly found in the brain. In contrast, Feulner et al. (2022) takes a more agnostic approach by effectively meta-training recurrent networks with feedback control of their output out-of-equilibrium and under non-stationary conditions. However, while the resulting networks reproduce key behavioural and neurophysiological findings from both humans and non-human primates, their relationship with gradient descent remains unclear.

3.2 Methods

3.2.1 Training stages

In this work, we used a modelling setup built to reproduce motor adaption experiments in monkeys (Perich et al., 2018) and similar to that used in Feulner et al. (2022) (see Appendix 3.6.6 for key differences). We construct a model that mimics both the movement repertoire and the flexibility of the motor cortex by applying a two-stage training strategy consisting of a pretraining and a fine-tuning step.

Pretraining. To obtain stereotypical motor movement dynamics, we first train the recurrent networks with feedback control on a synthetic *instructed delay centre-out reaching* task using BPTT (Werbos, 1990). Here, the task is to produce a broad set of two-dimensional reaching velocity trajectories of varying lengths, following an instructed delay phase (see Section 3.6.1 for more details). A visualization of a sub-task with 8 equidistant targets can be seen in Figure 3.1a-b.

Fine-tuning. To probe the recurrent network’s ability to adapt to task perturbations following initial pretraining, we replicate a classic visuomotor (VR) perturbation paradigm (Perich et al., 2018; Krakauer et al., 2000) in which the model output is rotated around the centre by a certain fixed angle. Here, motivated by rapid adaptations to such perturbations in the experimental studies (Perich et al., 2018), learning occurs through a local, biologically plausible learning rule.

3.2.2 Network architecture

We train the recurrent neural networks with a single recurrent hidden layer whose activity is described by the dynamics in Equation 3.1. Here, a hidden layer unit h_j receives the stimulus signal $s(t)$, the previous network activity post-nonlinearity $\phi(h_i)$ and the positional error ϵ_i from the previous time step i.e. $(t - 1)$.

$$\tau \dot{h}_j = \left(-h_j + \sum_i W_{ji}^{in} s_i + \sum_i W_{ji}^h \phi(h_i) + \sum_k W_{ji}^{fb} \epsilon_i + b_j^h \right) \quad (3.1)$$

In Equation 3.1 ϕ represents the ReLU function, and $\mathbf{W}^{fb}$ are the backward projecting feedback weights. The 2D output velocity $\mathbf{v}$ of the recurrent network is obtained through a linear readout of its hidden layer activity post-nonlinearity:

$$v_k = W_{kj}^o a_j + b_k^o \quad (3.2)$$

where $a_j = \phi(h_j)$.

3.2.3 Pretraining procedure

We pretrain the network using BPTT with mean squared error (MSE) over the integrated position space using the Adam optimizer (Kingma and Ba, 2014). To encourage the networks to learn to use the feedback layer to correct their output online, we use randomly distributed velocity perturbations (see Section 3.6.3 for more details). The 2D error feedback signal ϵ_k projected back to the network with feedback weights W^{fb} is the derivative of the mean squared error of the integrated position $p_k^t = p_k^{t-1} + dt \cdot v_k^t$, i.e. $\epsilon_k^t = p_k^{t*} - p_k^t$, where p_k^{t*} is the task target.

3.2.4 Adaptation and feedback-driven plasticity rule

To assess how the feedback integration may enable biologically plausible learning in the hidden recurrent layer during persistent task perturbation, we implement a previously proposed, temporally and spatially local learning rule, Random Feedback Local Online (RFLO) learning (Murray, 2019).

$$\tau \dot{B}_{ji} = \left(-B_{ji} + \phi'(h_j) a_i^{prev} \right) \quad (3.3)$$

$$\dot{W}_{ji}^h = \eta_2 W_{jk}^{fb} \epsilon_k B_{ji} \quad (3.4)$$

where η_2 is the adaptation learning rate and a_i^{prev} is hidden activity from the previous timestep i.e. $(t - 1)$. Thus, the change in weights is proportional to the eligibility trace B_{ji} and the feedback signal received $W_{jk}^{fb} \epsilon_k$. Note that the same weight matrix $\mathbf{W}^{fb}$ is used for both control and learning. For further biological realism, the learning batch size is fixed to 1 and the weight updates are applied online i.e. after every presented example.

3.3 Results

3.3.1 Recurrent neural networks with feedback control adapt well to task perturbation using feedback control and local learning rules

First, we pre-train the networks with feedback control on the *instructed delay center-out-reaching* task, where the networks need to produce 2D sigmoidal velocity profiles based on the target and the *go* signal input given (see Methods for details). Once trained, we test their ability to adapt to task perturbations using the visuo-motor rotation (VR) paradigm. Here, we rotate the target positions by 30° around the centre axis, creating a mismatch between the learned relationship between the input and the target. As previously shown in Feulner et al. (2022), the networks with feedback control show online adaptation to such a perturbation during ongoing movement in the first trial (Figure 3.1B). With persistent perturbation, the networks with feedback control show further improvement in the task performance, where the adapted network readout is independent of feedback control (Figure 3.1C).

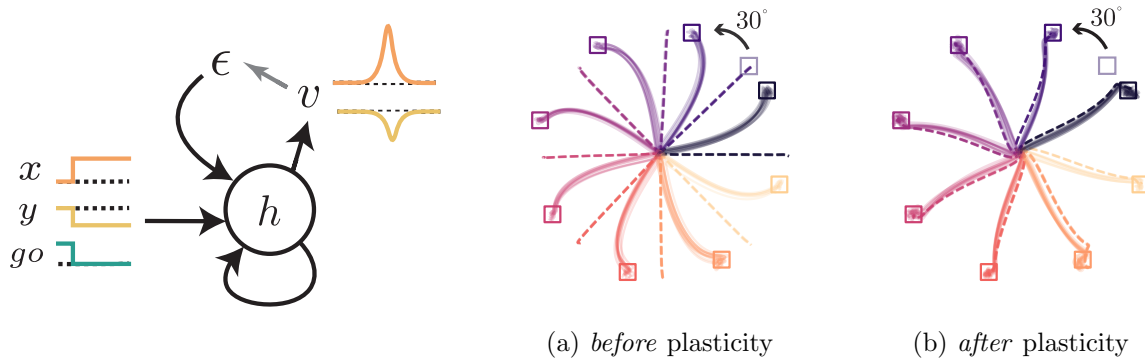


Figure 3.1: **Recurrent neural networks with feedback control feature rapid adaptation to acute task perturbation and can learn the persistent task perturbation using a local learning rule.** A recurrent neural network with feedback control is first pre-trained on produce a 2D sigmoidal velocity profile based on a target and *go* cue input input, and then tested on a 30° target rotation. (a) The pre-trained network adapts to the perturbation during the first trial with feedback control (full-line). The dashed line denotes the same network without feedback control. (b) The network further improves its performance during persistent perturbation with feedback-driven local learning (Feulner et al., 2022)

3.3.2 Feedback control approximates the true gradient w.r.t. networks activations during task perturbation

What happens during task perturbation in the networks with feedback control that allows for such rapid adaptation? In Figure 3.2a, we show in networks with feedback control, the feedback contribution to the overall network output increases during perturbation (see 3.6.7 for full analysis details). This suggests that during task perturbation, the network activity is increasingly driven by the feedback signal. The average magnitude of this drive is proportional to the movement perturbation magnitude (Figure 3.2b), as the feedback drive is used to correct the recurrent network activity during perturbed trials. To further understand the mechanism behind this rapid movement correction, we investigate the alignment of the online feedback signal $W^{fb}\epsilon^{t-1}$ with the true first-order gradient w.r.t. the network activations. We show that during predicted movement (between $t = 50$ and $t = 100$), the online feedback signal injected into the ongoing network dynamics approximates well the optimal, global trial gradient (Figure 3.2c). This adaptive behaviour is not typically observed in vanilla recurrent neural networks, where any performance feedback is “locked” i.e. not available to the network activity until the end of the whole trial, and even then only indirectly so through an update in the network weights. Thus, with feedback control, the feedback is “unlocked” and corrects the network activity in real-time.

3.3.3 Feedback control *and* local, feedback-driven learning enable rapid network adaptation during persistent task perturbation

We next focus on network adaptation during persistent perturbation with feedback-driven plasticity *only* in the recurrent layer. Here, plasticity is governed by the temporally and spatially local learning rule that uses the learned feedback weights to project error from the readout layer back to the recurrent layer (RFLO with control, i.e. RFLO+c, see Section 3.2.4). Such plasticity further improves the network performance during adaptation (Figure 3.3A, Feulner et al. (2022)). Thus, the network becomes independent of feedback control after a limited number of

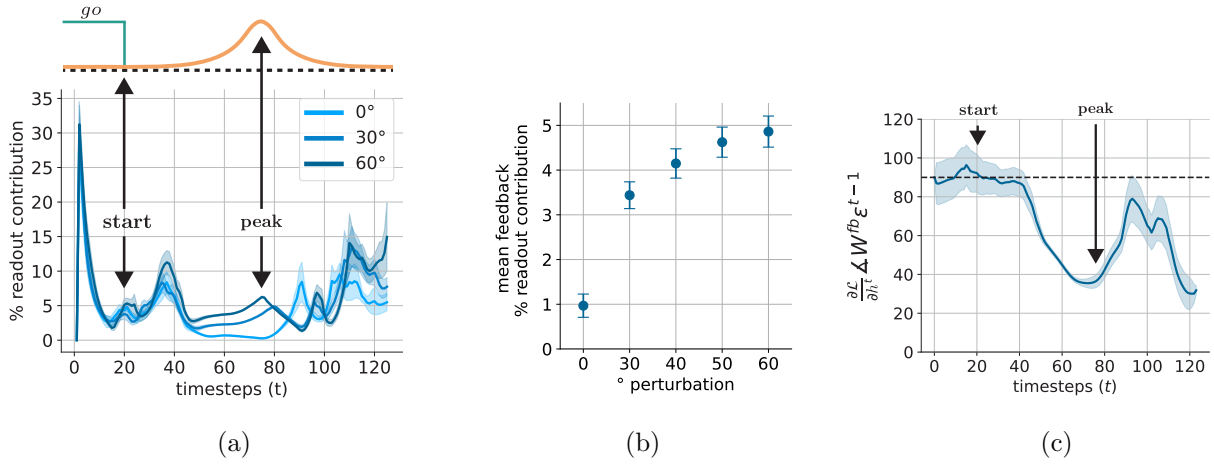


Figure 3.2: **Feedback control modulates rapid adaptation during task perturbation through approximate inference learning in the activity space.** (a) Increasing task perturbations degree ($0 \rightarrow 30 \rightarrow 60$) increases the relative feedback contribution to the network readout during a single task trial, which included a complete reaching movement. **start** denotes the start of the movement, and **peak** denotes the velocity peak. The shaded region denotes the standard error of the mean over 20 network seeds and all 8 targets. (b) The average magnitude of the feedback drive is proportional to the movement perturbation magnitude. (c) The cosine similarity of the online feedback signal injected into the ongoing network dynamics and the optimal, global trial gradient w.r.t. activations during a single task trial. The shaded region denotes the standard error of the mean over 20 network seeds and all 8 targets.

trials. Surprisingly, using this simple local rule with minor perturbations, networks learning using local learning rules show a performance similar to those trained with online BPTT (Figure 3.3b), with (RFLO +c) or without (RFLO) feedback control. However, with larger perturbations, networks trained using a local learning rule without feedback control show an increasing performance gap compared to those trained with feedback control and BPTT (Figure 3.3b).

3.3.4 Feedback control enables accurate recurrent weight adaptation during task perturbation

What make online recurrent network adaptation so hard in the first place, and how might feedback control help to guide learning? To answer this question, we outline the computational requirements of the true online gradient w.r.t. the network activations and the feedback-driven local learning rule (Equation 3.4) using the notation from Marschall et al. (2020). Using online BPTT (Werbos, 1990), i.e. Real-Time Recurrent Learning (RTRL) (Williams and Zipser, 1989), one can decompose the total loss $\mathcal{L}_t$ w.r.t. parameters $\mathbf{W}$ using the chain rule as follows:

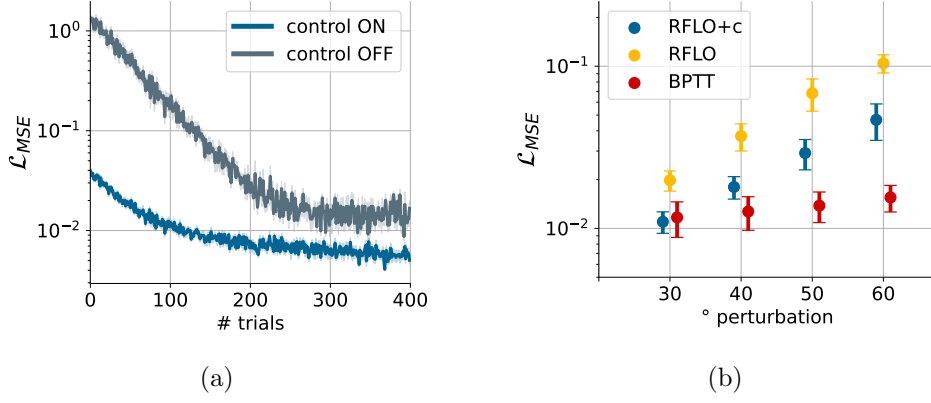


Figure 3.3: **Feedback control aids local learning in the recurrent network layer.** (a) Train loss as a function of trial number during persistent perturbation with feedback-driven local learning. Grey line denotes the tested performance of the same network without feedback control, showing that the network relies less on feedback control with learning. The shaded regions denote the standard deviation over 10 random network seeds. (b) The final performance (after 1000 trials) of recurrent networks trained with feedback-driven local learning with (blue) or without (yellow) feedback control during learning with varying degrees of persistent perturbation. The points in red denote the performance of networks trained with online BPTT. Note that all the performances reported here are in the setting without feedback control at test-time, irrespective of if feedback control was used during training. The error bars denote the standard deviation over 10 random network seeds.

$$\sum_t \frac{\partial \mathcal{L}_t}{\partial \mathbf{W}} = \sum_t \frac{\partial \mathcal{L}_t}{\partial \mathbf{a}_t} \frac{\partial \mathbf{a}_t}{\partial \mathbf{W}} \quad (3.5)$$

where $\frac{\partial \mathcal{L}_t}{\partial \mathbf{a}_t} = \bar{\mathbf{c}}_t$ is the credit assignment vector and $\frac{\partial \mathbf{a}_t}{\partial \mathbf{W}} = \mathbf{M}_t$ is the influence matrix (Marschall et al., 2020). Using RTRL, the credit assignment vector $\bar{\mathbf{c}}$ is *immediate* as it is simply calculated by backpropagating the immediate error ϵ^t through the derivative of the output function $F_{\mathbf{W}}^o$ ¹. However, the influence matrix $\mathbf{M}$ requires the recursive computation of the past network states using the network Jacobian $\mathbf{J}_t = \partial \mathbf{a}_t / \partial \mathbf{a}_{t-1}$ (for details, see Marschall et al. (2020)).

In this work, similar to previous work (Feulner et al., 2022; Bellec et al., 2020; Murray, 2019), $\mathbf{M}_t$ is approximated with an *eligibility trace* B (see Equation 3.4). However, as shown in (Marschall et al., 2020), such a severe approximation of the Jacobian can fail to capture the true temporal dependencies, instead biasing the learning trajectory towards capturing only short-term dependencies present in the neural activity at the point of learning. Depending on the task and the recurrent matrix structure, this can lead to the capturing of spurious

¹In this work, $\bar{\mathbf{c}}_t$ is approximated with a pre-learned feedback matrix $\mathbf{W}^{fb}$, which through pre-training approximately aligns with the transpose of $\mathbf{W}^o$.

correlations and a significant generalization gap (Liu et al., 2022). However, in networks with feedback control during task perturbation, the network activities are increasingly driven by the feedback signal (Figure 3.2a), and thus less so by their past, recurrent state. We can quantify this further by calculating the norm of the Jacobian of the current network hidden activity w.r.t. past hidden activity with and without feedback control during task perturbation, which gives us an indication of current’s networks sensitivity to its *past* activity vs the *present*, immediate feedback error. This is shown in Figure 3.4a, where we see that there is a drop in sensitivity to the past activity in the networks with feedback control during increasing task perturbation. Thus, as the network relies less on the past and more on the present, and is therefore “decoupled”, we expect the Jacobian approximation in the eligibility trace to be more accurate.

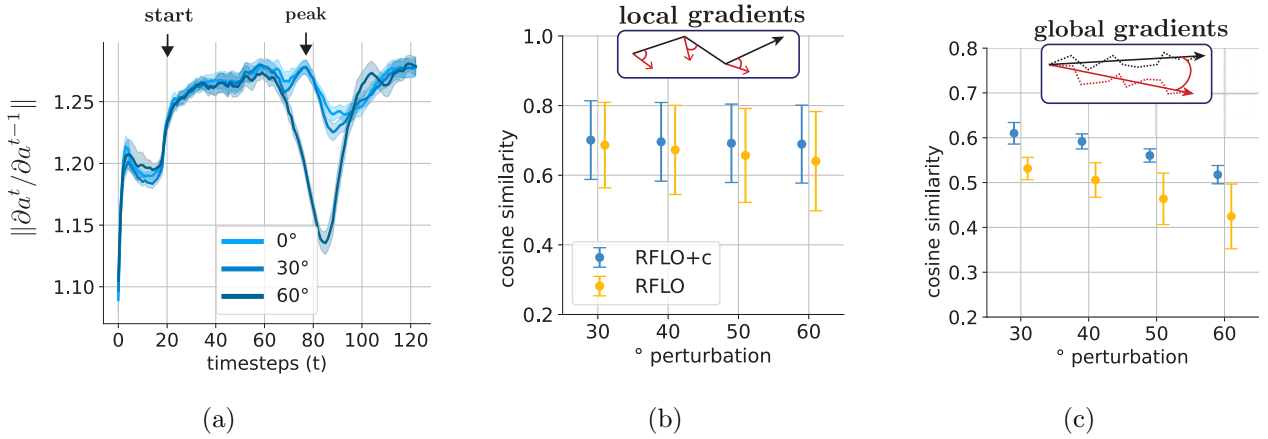


Figure 3.4: Feedback increases the accuracy of local learning. (a) The norm of the Jacobians of the network activities at time t w.r.t. to activities at a previous timestep during a single trial. The shaded regions denote the standard deviation over 10 random network seeds. (b) The mean local alignment of the feedback-driven local learning rule with the true local gradient during task adaptation. The error bars denote the standard deviation over 10 seeds. (c) The mean global alignment of the feedback-driven local learning rule with the true global gradient during task adaptation. The error bars denote the standard deviation over 10 random network seeds.

We empirically validate this hypothesis by comparing the gradients of the local learning rule with that of online BPTT during task adaptation, with and without feedback control. Note that we calculate both the local, i.e. at each learning step, and global, i.e. after all the learning steps, gradient alignments. We show that while there is only a minor difference in the local gradient alignments, this compounds at the global level, where the local learning rule in the networks

with feedback control aligns better with the true, global gradient than without feedback control (Figure 3.4b-c). This may explain the better performance of the networks with feedback control during task adaptation (Figure 3.3b), as the local learning rule is more accurate.

3.3.5 Feedback control enables efficient recurrent weight adaptation during task perturbation

To further understand the impact of feedback control on weight-based learning, we next focus on the relative magnitude of weight updates during task adaptation. Specifically, for each learning episode consisting of a limited number of trials, we can calculate the magnitude of final, global weight update $\|\Delta \mathbf{W}_{global}\|$, as well as the total combined magnitude of the online, local weight updates $\sum \|\Delta \mathbf{W}_{local}\|$. The ratio of the two quantities allows us to quantify the efficiency of the online weight updates during task adaptation. For instance, the less erratic the online weights updates are, the more efficient the learning trajectory, and our calculated ratio will be closer to 1. Despite the relatively minor difference in local gradient accuracy (Figure 3.4b), we show that our pre-defined “efficiency ratio” is significantly different between the network with and without feedback control during learning, at all tested perturbation magnitudes (Figure 3.5a). Thus, with feedback control, online weight updates are more efficient and less sensitive to the noise inherent to online learning using pseudo-gradients.

What could be the mechanism behind this marked increase in efficiency of the online weights updates with feedback control? In our networks, the feedback-controlled network activities are more constrained during task perturbation, where the network is initially more driven by the feedback signal that approximately aligns with their first-order gradient (see Figure 3.2c), irrespective of any weight based learning. This can be seen as a form of local adaptive inference in the activity space, or “prospective configuration” (Song et al., 2022), that can *guide* subsequent update steps in the weight space. Furthermore, the efficiency of learning is often linked to the network’s sensitivity to its higher-order gradients, which equips it with better navigation of pathological curvatures in the objective function (Martens et al., 2010). To assess this in our networks, we calculate the alignment of the online feedback control signal

injected into the network dynamics with the exact gradient used in the canonical 2^{nd} order optimization scheme, Newton’s method. Here, the gradient is defined as $(\mathbf{H} - \gamma \mathbf{I})^{-1} \mathbf{g}$, where $\mathbf{H}$ is the Hessian w.r.t. the network activations and γ is some small constant that re-conditions it. We show that during peaks of movement error, the feedback control signal injected into the ongoing network dynamics is also sensitive to their 2^{nd} order gradient (Figure 3.5b). Thus, feedback control also *implicitly* injects adaptive, second-order gradient information into the network dynamics, which may explain the magnitude of the increased efficiency observed in online weight updates during task adaptation.

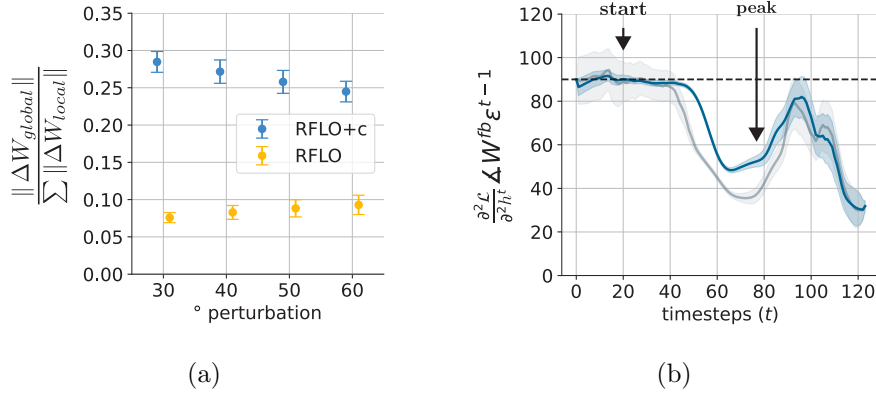


Figure 3.5: **Feedback increases the efficiency of local learning.** (a) The mean norm or efficiency ratio of the total, global weight update with that of the online, local weight update during task adaptation at various perturbation degrees. The error bars denote the standard deviation over 10 random network seeds. (b) The cosine similarity between the online feedback control signal injected into the ongoing network dynamics and the second-order gradient (calculated using the Newton method) w.r.t. activations (blue) during a single trial. The grey line denotes the first-order gradient as in Figure 3.2c. The shaded region denotes the standard error of the mean over 10 network seeds and all 8 targets.

3.4 Discussion

Summary. Although feedback connections are known to influence forward network activity in biological networks, the interaction of such feedback-driven activity for control with weight-based learning is less well understood. In this work, we investigated the mechanistic properties of recurrent neural networks trained with online feedback control and tested on task perturbation, which can also be seen as task modification, originally proposed in (Feulner et al., 2022). We show that the online adaptation of such networks during task perturbations is due to (1) the increased feedback drive of the network and (2) the approximate alignment of the feedback

signal with the true first and second-order gradient w.r.t. the network activations. Furthermore, we show that the consequence of such control is that the temporally and spatially local weight updates during task adaptation are both more accurate, robust and efficient.

Feedback control and learning. Our results are in line with recent theoretical work showing that injecting control signals into network activity may aid local, biologically plausible learning in systems at equilibrium (Meulemans et al., 2022b). Here, we build on this work by showing that such strict equilibrium conditions can be relaxed in the presence of learned online feedback control during an ongoing network trial. Moreover, we show that the feedback control integration in the activity space also *implicitly* injects adaptive, second-order gradient information into the network dynamics, which may explain the magnitude of the increased efficiency observed in online weight updates during task adaptation. This is also in line with recent theoretical work (Innocenti et al., 2023) linking (predictive-coding) network inference with the second-order, trust region optimization methods (Conn et al., 2000; Dauphin et al., 2014; Yuan, 2015). Here, unlike in standard line-search methods, a “safe” region in the loss landscape is first determined, before the actual weight updates within this region are calculated and applied. As in Innocenti et al. (2023), we show that in our networks this “safe” region could be implicitly calculated through adaptive, feedback-driven inference in the activity space. Importantly, in our networks this is done without an additional relaxation phase nor any explicit, second-order gradient calculation, and thus without the additional, often prohibitive, computational cost associated with such methods compared to a standard forward and backward network pass in BPTT.

Limitations & future work. In this work, we demonstrate the benefits of feedback control on a simple, biologically relevant task in single layer RNN. Further work is needed to investigate the generality of our findings to a wider range of tasks and network architectures. Moreover, we remain agnostic to the exact biological implementation of the feedback signal, and how it could be implemented in the brain (probably via cerebellum, see Pemberton et al. (2022)). Finally, the feedback form used in this work is a simple, linear projection of the signed error

signal back to the network, and we anticipate that this may not hold in many brain regions. We believe it would be interesting both for computational neuroscience *and* machine learning to investigate the impact of less explicit feedback mechanism. Thus, extending our results to learning conditions with increased generality is an important direction for future work.

3.5 Impact statement

This paper presents work whose goal is to advance our understanding of fundamental principles of biological motor control. We recognize that this could have far-reaching implications in the future, particularly in the fields of robotics and prosthetics. Beyond these applications, this research may also enhance our understanding of motor disorders, potentially leading to improved rehabilitation strategies. Careful ethical considerations throughout our research are vital to proactively navigate the benefits and potential challenges these advancements may introduce to society.

3.6 Appendix

3.6.1 Detailed Methods: Task

To create a synthetic *instructed delay centre-out reaching* task as experimentally performed in monkeys (Perich et al., 2018), we set both the trajectory start ($p^{start} \in \mathbb{R}^2$) and end point ($p^{end} \in \mathbb{R}^2$) by randomly sampling each coordinate independently from a uniform distribution $U(-6, 6)$ and interpolate the corresponding velocity trajectories $[v_x, v_y]$ with a sigmoid function parametrized by $\kappa = 10/s$ (Equation 3.6).

$$f(t) = \frac{1}{1 + \exp(-t\kappa)} \quad (3.6)$$

This creates natural, bell-shaped reaching velocity profiles. Each trial lasts ≈ 1.25 s (125 steps) and features a (fixed) instructed delay period t^{go} of 0.2 s, included in the network as a Heaviside step function that is non-zero during the delay period. Thus, the input and output task dimensionalities are 3 and 2 respectively. The radius of the circle for the *center-out-reach* sub-task used in visualisations is set to 5 cm.

3.6.2 Detailed Methods: Network

Here, we use recurrent neural networks with 400 hidden units. The weights ($\mathbf{W}^{in}, \mathbf{W}^h, \mathbf{W}^o, \mathbf{W}^{fb}$) and biases (b^h, b^o) of the networks are initialized, following the strategy from He et al. (2015), from a uniform distribution $U(-1/\sqrt{l}, 1/\sqrt{l})$ where l is either the number of layer parameters (for W^{in}, W^h, W^{fb}, b^h) or the dimensionality of the output (for W^o, b^o). The hidden network activity x is integrated using Equation 1, with $dt = 10$ ms and $\tau = 50$ ms.

3.6.3 Detailed Methods: Pretraining

We initially train the networks using the Adam optimizer (Kingma and Ba, 2014) with learning rate $\eta_1 = 0.001$ ($\beta_1 = 0.9, \beta_2 = 0.999$), batch size of 256 and L2 regularization for network parameters ($\beta = 1e^{-3}$) and activations ($\gamma = 2e^{-3}$). The total norm of the gradients is clipped to 0.2. Here, we eliminate the separate parameter training phases used in Feulner et al. (2022) and keep all network parameters plastic throughout the initial training. We also perturb the network output during the entire pretraining stage by adding brief (0.1 s) velocity pulses with an amplitude of 10 cm/s at various movement periods in 25% of the trials.

3.6.4 Detailed Methods: Fine-tuning

During online adaptation to perturbation, we use the learning rate $\eta_2 = 1e^{-5}$.

3.6.5 Detailed Methods: Computational requirements

We ran all our experiments using NVIDIA GeForce GPUs (RTX 2080 Ti).

3.6.6 Key differences from Feulner et al. (2022)

- **Task Simplification:** We simplify the task by fixing the *go* signal timing, facilitating a clearer analysis of within-trial adaptation mechanisms.
- **Curriculum Streamlining:** We use a single-phase training curriculum instead of the three-phase approach, again focusing on the core mechanisms of interest.
- **Pre-training batch size increase:** We use a larger batch size than in Feulner et al. (2022) (20 vs 246), for faster convergence of networks with and without feedback.
- **Online feedback:** We provide truly online feedback, removing the biologically-motivated feedback delay used in Feulner et al. (2022).
- **Shorter Individual Trials:** We shorten the individual trial lengths (300 to 125) to improve the distribution of target values - in Feulner et al. (2022), most of the target values are set to 0 due to the steepness of the velocity sigmoid.
- **Local Learning Rule:** We improve the biological plausibility of the learning rule by making it fully online (instead of accumulated) with a continuous eligibility trace (see Equation 3.4). We also add $\phi'(a_j)$ to the eligibility trace, which makes it equivalent to that of RFLO learning (Murray, 2019) and increases its stability with larger perturbations (see Figure 3.6 and 3.7).

3.6.7 Detailed Methods: Analysis details

To quantify the percentage readout contribution of the immediate feedback control signal in Figure 3.2a, we revisit Equations 3.1 and 3.2 that govern the network hidden and readout

dynamics, respectively. Combining the two into a single equation, we get:

$$v_k = \sum_j W_{kj}^o a_j + b_k^o \quad (3.7)$$

$$= \sum_j W_{kj}^o \phi(\dot{h}_j) + b_k^o \quad (3.8)$$

$$= \sum_j W_{kj}^o \phi \frac{dt}{\tau} \left(-h_j + \sum_i W_{ji}^{in} s_i + \sum_i W_{ji}^h \phi(h_i) + \sum_i W_{ji}^{fb} \epsilon_i + b_j^h \right) + b_k^o \quad (3.9)$$

Then, the percentage of the immediate feedback control signal % contribution to the network output is easily calculated as:

$$\text{percentage feedback contribution} = \sum_j [h_j > 0] W_{kj}^o \left(\frac{dt}{\tau} \sum_i W_{ji}^{fb} \epsilon_i \right) \times \frac{100}{\sum_j W_{kj}^o a_j} \quad (3.10)$$

Note that for Figure 3.2a we report an average value over k .

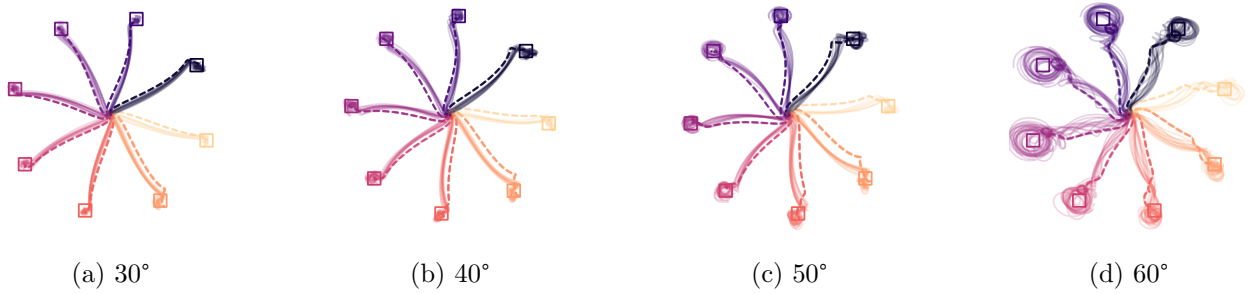


Figure 3.6: Performance of networks trained using the local rule from Feulner et al. (2022) with increasing degrees of perturbation.

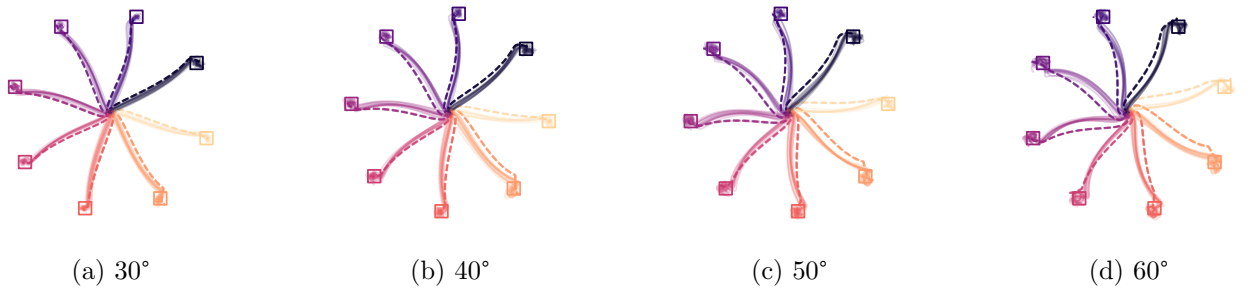


Figure 3.7: Performance of networks trained using RFLO learning from Murray (2019) with increasing degrees of perturbation.

Chapter 4

Meta learning feedback plasticity rules

This chapter delves into the mechanisms of learning biologically plausible learning in cortical multi-layer networks. Specifically, it addresses the "weight transport" problem by exploring the space of local and effective feedback learning rules using meta-learning and a flexible student-teacher distillation task. It demonstrates that meta learning using gradient descent can be successfully applied to learning the feedback plasticity rules, as well as the feedback pathway architecture.

To be consistent with previous literature, in this chapter the network notation from Nøkland (2016) is used. Specifically, network hidden activity before the non-linearity is denoted by $\mathbf{a}$ and after non-linearity as $\mathbf{h}$. Note that this is in contrast with Chapter 3.

A version of this work was presented at Cosyne 2022 and is currently unpublished.

Kaleb and Clopath (2022). Meta-learning biologically plausible feedback learning rules. COSYNE 2022 19 March / III-010 in Lisbon, Portugal. [\[link\]](#)

Meta-learning biologically plausible feedback learning rules

Klara Kaleb¹, Claudia Clopath¹

¹Imperial College London, Department of Bioengineering, London, UK

Abstract

Learning in cortical multi-layer networks is thought to occur through spatially and temporally local synaptic plasticity aided by ubiquitous feedback connections from upper layers. However, the known biologically plausible learning rules for such networks have yet to reach the high performance of machine learning algorithms, such as backpropagation, on complex tasks that the brain can solve. It has been shown that one of the major obstacles to backpropagation-like algorithm implementation in the brain, the *weight transport* problem, can be overcome with initially random feedback connections undergoing local synaptic plasticity. Nevertheless, the performance gap remains. Here, we use the meta-learning framework to further explore the space of effective biologically plausible feedback learning rules. More precisely, we meta-learn a parametric function defining the learning rule using gradient descent on a task defined using the flexible student-teacher setup. Moreover, we meta-learn both the feedback learning terms and the architecture of the feedback pathways itself. We show that our meta-learned architecture and learning rule robustly close the gap between backpropagation and biologically plausible learning by aligning the pseudo-gradients to the true gradients after a limited number of learning steps. Furthermore, we show that the solutions found by our learning rule are nearer the initial network parametrization than those found by feedback alignment, which suggests that the meta-learned learning rule is learning more efficiently. In summary, we show that meta-learning of learning rules can also be applied to the plasticity of feedback network connections and provide insights into how multi-layer learning may operate in the brain.

4.1 Introduction

The remarkable success of deep learning using backpropagation (Rumelhart et al., 1986) in artificial networks has raised the question of whether the brain employs similar learning mechanisms. However, biological systems like the brain operate under constraints that challenge the direct applicability of backpropagation. Backpropagation’s biological implausibility stems from several critical discrepancies with the known workings of the brain (Crick, 1989; Mazzoni et al., 1991; O’Reilly, 1996; Chinta and Tweed, 2012; Bengio et al., 2015; Lillicrap et al., 2020). These include the separation of error gradients from neuronal activations, differences between forward and backward pass computations, and most notably, the *weight transport* problem (Grossberg, 1987). The *weight transport* problem arises from the requirement of symmetric weights for feedforward and feedback connections, a condition not found in the brain. This fundamental mismatch has led to the exploration of alternative, biologically plausible learning mechanisms that can better align with the complexity of neural circuitry.

Several solutions have been proposed to address the *weight transport* problem in the brain. Notably, it has been demonstrated that learning can occur with an independent set of randomly initialized feedback weights that connect each subsequent network layer (Lillicrap et al., 2016) or provide direct connections from the network output layer to all hidden layers (Nøkland, 2016). Here, the plastic forward weights gradually align with the fixed random feedback weights, enabling effective learning. However, such a learning approach may require network over-parametrization for degeneracy breaking in the solution space (Refinetti et al., 2021), and may struggle with small datasets and deep networks (Liao et al., 2016). Furthermore, the learning rules may require large batch sizes to work effectively (Liao et al., 2016), which is not consistent with the brain’s remarkable ability to learn fully *online*, i.e. with a batch size of 1 for a single epoch (Goodfellow et al., 2016).

Several studies have explored potential mechanisms that could enhance the alignment of the feedback weights with the forward weights, and thus improve learning in more challenging settings. These approaches include sign symmetry (Liao et al., 2016) and independent plasticity mechanisms in the feedback weights (Akrouit et al., 2019; Kunin et al., 2020; Kolen and Pollack,

1994). Notably, Kolen and Pollack (1994) showed that updating feedback weights using the transpose of the forward weight learning rule, coupled with weight decay on both sets of weights, leads to their convergence. More recently, such feedback learning approaches have been shown to scale to complex tasks like ImageNet using large, residual neural networks (Akrouf et al., 2019; Kunin et al., 2020). While such feedback-learning approaches have made significant strides, their ability to match the brain’s online learning capabilities and robustness remains an open question.

Beyond manual learning rule derivation, meta-learning i.e. *learning to learn* (Schmidhuber, 1987; Hinton and Plaut, 1987; Bengio et al., 1990) offers an alternative approach to discovering effective biologically plausible learning. Here, a learning rule is defined as a parametric function, the meta-parameters of which are learned using an optimization algorithm of choice. This approach has previously shown promise in discovering effective forward learning rules for feedforward networks (Bengio et al., 1990, 2013; Jordan et al., 2021; Confavreux et al., 2020; Shervani-Tabar and Rosenbaum, 2023; Confavreux et al., 2024). However, whether meta-learning can work on the complex problem of feedback plasticity has yet to be shown.

In this work, we explore the potential of meta-learning to discover effective *and* local feedback-learning rules for multi-layer networks using a flexible knowledge distillation task. In addition to meta-learning the learning rules, we also meta-learn the network feedback architecture. Our main contributions are:

- The synthetic student-teacher knowledge distillation setup is a useful tool for varying learning task complexity and thus revealing performance differences between different learning algorithms 4.4.1.
- Meta-learning of feedback plasticity under biological constraints can successfully identify a limited of learning rule parameters *and* the optimal feedback network architecture 4.4.2.
- The meta-learned rule aligns the pseudo-gradients within a limited number of learning steps, improving the learning efficiency compared to direct feedback alignment 4.4.3.

4.2 Background

4.2.1 Learning in Multi-layer Feedforward networks

To illustrate the biological difficulty with learning in a deep neural network, consider a network with L layers. Let $\mathbf{x} \in \mathbb{R}^N$ be the input vector, $\mathbf{W}_l$ be the weight matrix at layer l , ϕ be the activation function, and $\mathbf{y}$ be the output vector. The standard forward pass can then be represented by the following equations:

$$\mathbf{h}_0 \triangleq \mathbf{x} \quad (4.1)$$

$$\mathbf{a}_l = \mathbf{W}_l \mathbf{h}_{l-1}, \forall 1 \leq l < L \quad (4.2)$$

$$\mathbf{h}_l = \phi(\mathbf{a}_l), \forall 1 \leq l < L \quad (4.3)$$

$$\mathbf{y} = \mathbf{W}_L \mathbf{h}_{L-1} \quad (4.4)$$

In order to learn the weights of a neural network, we typically use backpropagation to compute the gradients of the loss function $\mathcal{L}$ with respect to each weight $\mathbf{W}_l$. This process involves calculating the error gradients for each layer and propagating them backward through the network. Let's define the error gradient at layer l as δ_l , which is the gradient of the loss with respect to the activations $\mathbf{h}_l$ at layer l :

$$\delta_l \triangleq \frac{\partial \mathcal{L}}{\partial \mathbf{h}_l} \quad (4.5)$$

To propagate the error backward from layer l to layer $l-1$, we need to compute δ_{l-1} , the gradient of the loss with respect to the activations at layer $l-1$. This is done using the chain rule:

$$\delta_{l-1} = \frac{\partial \mathcal{L}}{\partial \mathbf{h}_{l-1}} = \frac{\partial \mathcal{L}}{\partial \mathbf{h}_l} \underbrace{\frac{\partial \mathbf{h}_l}{\partial \mathbf{a}_l}}_{\phi'(\mathbf{a}_l)} \underbrace{\frac{\partial \mathbf{a}_l}{\partial \mathbf{h}_{l-1}}}_{\mathbf{W}_l} \quad (4.6)$$

$$= \delta_l \cdot \phi'(\mathbf{a}_l) \cdot \mathbf{W}_l \quad (4.7)$$

Here, $\phi'(\mathbf{a}_l)$ represents the derivative of the activation function, and $\mathbf{W}_l$ is the weight matrix of the l -th layer.

In biological neural networks, there is no clear mechanism for ensuring that the weights used for propagating errors backward are the same weights used in the forward pass (Lillicrap and Santoro, 2019). Feedback alignment (FA) (Lillicrap et al., 2016) addresses this problem by using a separate set of fixed random weights $\mathbf{B}_l$ for error propagation to the hidden layers. In this case, we substitute δ_l above with δ_l^R , defined as:

$$\delta_{l-1}^R = \mathbf{B}_l (\phi'(\mathbf{a}_l) \odot \delta_l) \quad (4.8)$$

where $B_l \in \mathbb{R}^{N_l \times N_{l+1}}$. We refer to this as *regular* feedback alignment. In an alternative formulation, the error is directly transported from the output layer L to each hidden layer l (Nøkland, 2016). In this case, we define δ_l^D as:

$$\delta_{l-1}^D = \mathbf{D}_l \epsilon_L \quad (4.9)$$

where $D_l \in \mathbb{R}^{N_l \times N_L}$. This is referred to as *direct* feedback alignment. Should the network contain both types of feedback weight, we can easily define a parameter α that determines the relative contributions to δ_l from $\mathbf{B}_l$ and $\mathbf{D}_l$ as:

$$\delta_l = (1 - \alpha)\delta_l^D + \alpha\delta_l^R \quad (4.10)$$

Moreover, both $\mathbf{D}$ and $\mathbf{B}$ may undergo plasticity themselves, and thus aid learning in the forward weights $\mathbf{W}$ (Akrouit et al., 2019; Kunin et al., 2020; Kolen and Pollack, 1994).

4.2.2 Meta-Learning of *Biologically Plausible* Learning Rules

Meta-learning is a subfield of machine learning that focuses on *learning to learn* (Schmidhuber, 1987; Hinton and Plaut, 1987; Bengio et al., 1990). Here, we focus on gradient based meta-learning. There are multiple flavours of gradient-based meta-learning, such as learning initialization, exemplified by MAML (Model-Agnostic Meta-Learning) (Finn et al., 2017; Lindsey and Litwin-Kumar, 2020), and learning optimizers, such as in (Andrychowicz et al., 2016; Metz et al., 2019, 2022; Shervani-Tabar and Rosenbaum, 2023). Our approach rests in the latter category, where we learn the feedback learning rule itself ¹. Here, we apply candidate learning rules to a learning task, evaluate their performance, and update the learning rules accordingly.

Meta-learning typically consists of an inner loop and an outer loop. In the inner loop, the learning rule is applied to update the network parameters over a number of learning steps. In the outer loop, the meta-parameters of the learning rule are updated based on the performance of the network on a validation set.

Mathematical Formulation of Meta-Learning

In meta-learning, the goal is to optimize a learning algorithm such that it performs well on a variety of tasks. Let's denote a learning rule as f_θ . The learning rule is applied within an inner loop on a training dataset $\mathcal{D}_{train} = \{x_i, y_i\}_{0 \leq i < N_s}$, resulting in an update of the network parameters $\mathbf{B}$ and $\mathbf{D}$, starting with some initial parameters $\mathbf{B}^{(0)}$ and $\mathbf{D}^{(0)}$, randomly initialised at the start of each inner loop. Formally, we can express this as:

$$\mathbf{B}^{(0)} \sim p_0 \tag{4.11}$$

$$\mathbf{B}^{(i+1)} \leftarrow \mathbf{B}^{(i)} - \eta_B \cdot f\left(x_i, y_i, \mathbf{B}^{(i)}; \boldsymbol{\theta}_B^{(t)}\right) \quad \forall 0 \leq i < N_S \tag{4.12}$$

$$\mathbf{D}^{(0)} \sim p_0 \tag{4.13}$$

$$\mathbf{D}^{(i+1)} \leftarrow \mathbf{D}^{(i)} - \eta_D \cdot f\left(x_i, y_i, \mathbf{D}^{(i)}; \boldsymbol{\theta}_D^{(t)}\right) \quad \forall 0 \leq i < N_S \tag{4.14}$$

¹Note that we assume the forward weights are learning using 4.8 and/or 4.9

where η_B and η_D are the learning rates. The performance of the learning rule is evaluated on a validation set $\mathcal{D}_{val}$, and the meta-objective is to minimize this validation loss:

$$\mathcal{L}_{meta}(\boldsymbol{\theta}) = \mathcal{L}(\mathbf{B}^{(N_S)}, \mathbf{D}^{(N_S)}, \mathcal{D}_{val}) \quad (4.15)$$

To update $\boldsymbol{\theta}$, we use a gradient-based method:

$$\boldsymbol{\theta}_B^{(t+1)} \leftarrow \boldsymbol{\theta}_B^{(t)} - \eta_{meta} \cdot \nabla_{\boldsymbol{\theta}} \mathcal{L}_{meta}(\boldsymbol{\theta}_B^{(t)}) \quad (4.16)$$

$$\boldsymbol{\theta}_D^{(t+1)} \leftarrow \boldsymbol{\theta}_D^{(t)} - \eta_{meta} \cdot \nabla_{\boldsymbol{\theta}} \mathcal{L}_{meta}(\boldsymbol{\theta}_D^{(t)}) \quad (4.17)$$

where η_{meta} is meta-learning learning rate.

4.3 Methods

4.3.1 The learning task

In this work, we design a synthetic regression task using the flexible student-teacher framework (Gardner and Derrida, 1989; Seung et al., 1992; Zdeborová and Krzakala, 2016) on a knowledge distillation task (Hinton et al., 2015). Here, a randomly initialized and fixed *teacher* network defines the target function for another, plastic *student* network to learn. The output of the *teacher* network, denoted as $\mathbf{y}_T$, is given by:

$$\mathbf{y}_T = f_T(\mathbf{x}; \mathbf{W}_T) \quad (4.18)$$

where f_T represents the *teacher* network function, $\mathbf{x}$ represents the input data and $\mathbf{W}_T$ are the fixed weights of the teacher network. The *student* network aims to learn the mapping defined by the teacher network. The output of the student network, denoted as $\mathbf{y}_S$, is given by:

$$\mathbf{y}_S = f_S(\mathbf{x}; \mathbf{W}_S) \quad (4.19)$$

where f_S represents the student network function, and $\mathbf{W}_S$ are the plastic weights of the student network. The objective is to minimize the difference between the outputs of the teacher and student networks. We define the loss function as the mean squared error (MSE) between $\mathbf{Y}_T$ and $\mathbf{Y}_S$:

$$\mathcal{L}(\mathbf{W}_S) = \frac{1}{d_{\text{out}}} \sum_{j=1}^{d_{\text{out}}} (f_T(\mathbf{x}; \mathbf{W}_T)_j - f_S(\mathbf{x}; \mathbf{W}_S)_j)^2 \quad (4.20)$$

where d_{out} is the number of output dimensions.

By keeping the teacher network depth fixed but varying the width, we can systematically modulate the task complexity for the student to learn from the teacher, as shown empirically in Figure 4.1a. Note that some care must be taken with this synthetic task design approach, as discussed in Appendix 4.6.1.

In this work, we focus on the underparametrized regime where the student network cannot simply memorize the teacher network. The teacher network has 4 layers of width 1024, and the student network has 3 layers of width 64. Both networks have 100 input and output units, and all their weights are initialized with Xavier normal initialization (Glorot and Bengio, 2010):

$$w_{ij} \sim \mathcal{N}\left(0, \frac{\text{gain} \cdot 2}{d_{\text{in}} + d_{\text{out}}}\right) \quad (4.21)$$

where $\text{gain} = 1.5$, and d_{in} (d_{out}) is the layer input (output) dimension. Both networks use the tanh activation function, and the input data is randomly sampled from a uniform distribution.

4.3.2 The learning rule formulation

In this work, we aim to discover effective local learning rules for the feedback weights of the student network. We constrain feedback plasticity to depend only on the local parameters of the layer, i.e., the feedback weight itself ($\mathbf{B}_l$ or $\mathbf{D}_l$), as well as the activations $\mathbf{a}$ and error ϵ at the current and previous layers ($l, l-1$). We define the feedback learning rule as $f(\theta)$, which is a linear combination of the local learning rule terms f_i weighted by a set of learned

meta-parameters θ shared across layers:

$$f(\theta) = \sum_{i=0}^N \theta_i \cdot f_i \quad \text{where} \quad f_i \in \{f_1, f_2, \dots, f_N\} \quad (4.22)$$

Precisely, if N_l represents the number of units on layer l , then the feedback weights $\mathbf{B} \in \mathbb{R}^{N_l \times N_{l-1}}$, the activations $\mathbf{a}_l \in \mathbb{R}^{N_l}$, and the error $\boldsymbol{\epsilon}_l \in \mathbb{R}^{N_l}$ are used to construct the $f_i \in \mathbb{R}^{N_l \times N_{l-1}}$. For the individual learning rule terms, we consider the following:

$$\begin{aligned} f_0 &= \mathbf{a}_l \boldsymbol{\epsilon}_{l-1}^\top & f_7 &= \mathbf{B}_l \odot (\mathbf{a}_l \mathbf{a}_{l-1}^\top) \odot (\boldsymbol{\epsilon}_l \boldsymbol{\epsilon}_{l-1}^\top) \\ f_1 &= \boldsymbol{\epsilon}_l \mathbf{a}_{l-1}^\top & f_8 &= \mathbf{a}_l \mathbf{a}_{l-1}^\top \\ f_2 &= \boldsymbol{\epsilon}_l \boldsymbol{\epsilon}_{l-1}^\top & f_9 &= \mathbf{B}_l \odot (\mathbf{1}_{N_l} \mathbf{a}_{l-1}^2{}^\top) \\ f_3 &= \mathbf{B}_l & f_{10} &= \mathbf{B}_l \odot (\mathbf{a}_l^2 \mathbf{1}_{N_{l-1}}^\top) \\ f_4 &= \boldsymbol{\epsilon}_l \mathbf{1}_{N_{l-1}}^\top & f_{11} &= \mathbf{B}_l \odot (\boldsymbol{\epsilon}_l^2 \mathbf{1}_{N_{l-1}}^\top) \\ f_5 &= (\mathbf{a}_l \mathbf{a}_{l-1}^\top) \odot (\mathbf{1}_{N_l} \boldsymbol{\epsilon}_{l-1}^\top) & f_{12} &= \mathbf{B}_l \odot (\mathbf{1}_{N_l} \boldsymbol{\epsilon}_{l-1}^2{}^\top) \\ f_6 &= \mathbf{B}_l \odot (\boldsymbol{\epsilon}_l^2 \mathbf{a}_{l-1}^2{}^\top) \end{aligned} \quad (4.23)$$

where $\odot$ is the element-wise Hadamard product. Note that this is similar to the learning rule term pool for the forward learning rule used in Shervani-Tabar and Rosenbaum (2023), with the key differences being (1) the inclusion of the feedback weights instead of forward weights, (2) the addition of a separate, purely Hebbian term f_8 , (3) strict locality constraint on the Oja-like regularization terms (f_9 and f_{10})², and finally (4) the incorporation of additional Oja-like regularization terms that include the error term ϵ (f_{11} and f_{12}) (Oja, 1982). It is important to keep in mind that, for feedback weights, the pre-synaptic and post-synaptic roles are implicitly reversed as they transport error signals backward through the network. Note that the learning rule formulation for $\mathbf{D}_l$ is the same, with $\mathbf{D}$ replacing $\mathbf{B}$ and the pre- and post-synaptic quantities updated accordingly.

²i.e. each synaptic weight update only depends on its local pre-synaptic and post-synaptic quantities, and not on any more global network signals such as $\mathbf{a}_l \mathbf{a}_l^\top$

4.3.3 Meta-learning setup

Our meta-learning setup consists of a stereotypical inner and outer loop as defined in 4.2.2. In the inner loop, we apply the learning rule to the network feedback weights for a number of learning steps ($N_s = 300$) with a fixed learning rate $\eta_B = \eta_D = 1.0$. The networks are initialized with both regular and direct feedback weights. Their relative contributions to backward error propagation and thus feedback pathway architecture is determined by an additional meta-learned parameter α , as defined in 4.10. The parameter α is restricted to the $[0, 1]$ value range by a sigmoid function, and initialized at 0.0. The two sets of feedback weights $\mathbf{B}$ and $\mathbf{D}$ have a separate learning rule formulation that is shared across all network layers. Thus, we meta-learn $2 \times 13 + 1 = 27$ parameters in total. The student networks learn fully online, i.e., one example at a time for a single epoch, and all weights are updated after each example presentation.

In the outer loop, we calculate the meta-gradient on a held-out validation loss $\mathcal{L}_{\text{val}}$ with respect to each meta-parameter and update them accordingly using the ADAM gradient descent optimizer (Kingma and Ba, 2014). We include an additional L1 regularization term weighted by γ over all θ to encourage sparse learning rule formulations. This process is repeated for a given number of outer meta-iterations, each featuring a shuffled order of data presentation and a random network initialization. Note that within each inner loop we train 2 randomly initialised student networks independently on 2 different tasks, and ensemble the meta-gradients.

4.4 Results

4.4.1 Underparametrized student-teacher setup reveals a performance gap between backpropagation and feedback alignment

To assess reliable credit assignment across deep network layers, a non-linear task is required that demands error propagation to deeper layers beyond just the readout layer plasticity. The task should be complex enough to differentiate learning rule performance, yet computationally

tractable within the meta-learning framework. First, the performance of student networks trained with different learning rules is evaluated on teacher networks of varying widths. The results show that student networks trained with backpropagation (BP) outperform those trained with feedback alignment (FA) across all teacher network sizes, as shown in Figure 4.1a. Notably, the performance gap between BP and FA increases with the teacher network size, indicating that FA struggles with increased task complexity. Additionally, the performance of student networks trained with direct feedback alignment (dFA) and the delta rule (DR), where only the readout layer is plastic, is compared to those trained with BP and FA. The findings reveal that dFA outperforms FA on all teacher network sizes but still lags behind BP, while DR performs the worst. This highlights the importance of error propagation to hidden layers for effective learning in non-linear tasks. Figure 4.1b illustrates the dynamics of the student network test loss as a function of the number of presented examples for a teacher network with 1024 hidden units, showing that the values reported in Figure 4.1a are at convergence.

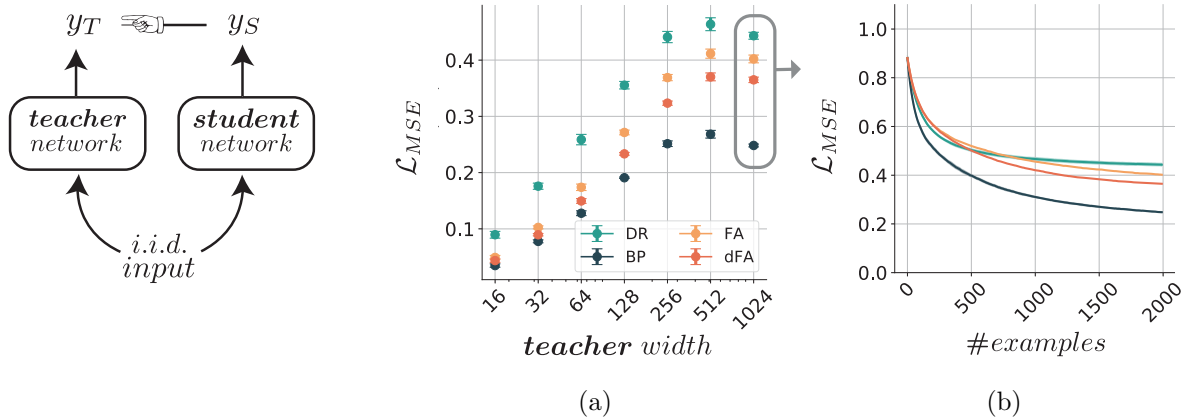


Figure 4.1: **A simple student-teacher setup reveals a performance gap between backpropagation and feedback alignment.** A feedforward *student* network is trained to approximate the output generated by a randomly initialised and fixed *teacher* network. Both networks receive the same i.i.d. input. (a) Test loss of a (3 hidden layer) student networks as a function of (4 hidden layer) teacher network width. The student networks have 64 hidden units per layer and are trained with either backpropagation (BP), feedback alignment (FA), direct feedback alignment (dFA) or delta rule (DR). With DR, only the readout layer is plastic, and no error propagation occurs to the deeper layers. Here, the error bars denote the standard deviation over 4 randomly initialized student *and* teacher networks. (b) Test loss of the same student network as a function of number of presented examples. Here, the learning task is a teacher network with a width of 1024 hidden units.

4.4.2 Meta-learning identifies a limited set of learning rule parameters for the direct feedback weights

The meta-learning framework is then applied to discover effective feedback learning rules for the student networks. Specifically, two sets of parameters are meta-learned: the learning rule parameters for the regular feedback weights and the direct feedback weights. Additionally, the proportion of the error coming from the regular feedback connections versus direct feedback connections, denoted as α , is also learned.

The α parameter converges to 0 (Figure 4.3b), indicating that the regular feedback weights are not used at all, making the initial learning rule equivalent to that of direct feedback alignment (dFA). Consequently, the feedback learning rule terms for regular feedback connections all converge to 0 (Figure 4.3c). However, a set of direct feedback learning rule terms converges to non-zero values, as shown in Figure 4.3d. This robust effect is observed across different random initializations. The terms selected are the following: $(+)\epsilon_l \mathbf{a}_{l-1}^\top(f_1)$, $(-)\mathbf{a}_l \mathbf{a}_{l-1}^\top(f_8)$ and $(+)\mathbf{B}_l \odot (\mathbf{1}_{N_l} \epsilon_{l-1}^2)^\top(f_{12})$.³

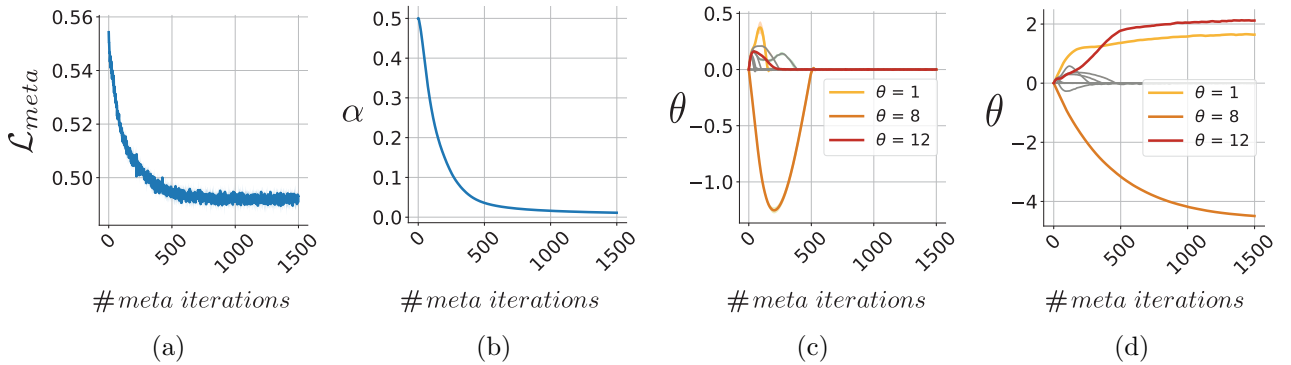


Figure 4.2: **Meta-learning identifies a limited set of learning rule parameters for the direct feedback weights.** (a) Meta loss as a function of meta iterations. (b) Parameter α , which denotes the proportion of the error coming from regular feedback connections, as a function of meta iterations. (c) Regular feedback weight parameters as a function of meta iterations. (d) Direct feedback weight parameters as a function of meta iterations.

³Note that these terms represent the *gradients*, so the update direction is *against* them.

4.4.3 Meta-learned learning rule closes the performance gap between backpropagation and direct feedback alignment and aligns the pseudo-gradients with a limited number of examples

By plugging the meta-learned learning rule into the student networks, we show that the performance gap between backpropagation (BP) and direct feedback alignment (dFA) is significantly reduced, as shown in Figure 4.3a. Moreover, our meta-learned rule avoids a common pathology in meta-learning, namely that of myopia, where the meta-learned optimizers fail to generalize beyond the constrained inner-loop learning window (Wu et al., 2018; Metz et al., 2019), as it generalizes beyond the 300-example window seen during meta-learning.

Furthermore, the pseudo gradients of the student networks obtain using the plastic direct feedback weights rapidly align with the true gradients with a limited number of examples, as shown in Figure 4.3b. This rapid alignment suggests that the meta-learned rule effectively approximates the optimal learning dynamics, enhancing the efficiency of the learning process. Additionally, the solutions found using the meta-learned rule are almost as efficient as those found using BP in terms of proximity to the initial parameters across all layers, as depicted in Figure 4.3c. In contrast, solutions for dFA diverge more significantly from the initial parameters, indicating less efficient learning in terms of total parameter changes required.

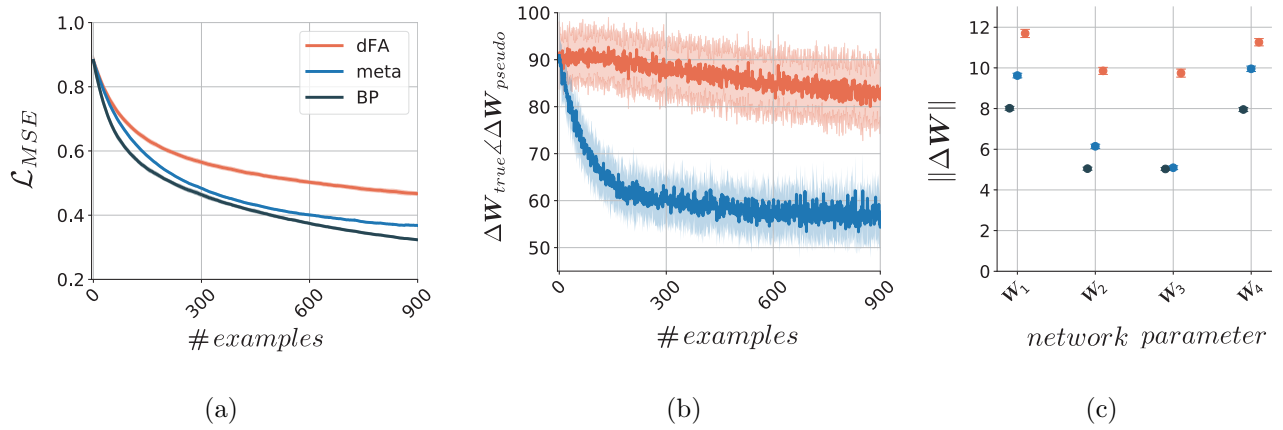


Figure 4.3: **Meta-learned learning rule closes the performance gap between backpropagation and direct feedback alignment and aligns the pseudo-gradients with a limited number of examples.** (a) Inner loss as a function of inner iterations for the meta-learned learning rule (meta, light blue), backpropagation (BP, dark blue) and direct feedback alignment (dFA, orange). (b) Inner gradients as a function of inner iterations for the meta-learned learning rule and direct feedback alignment. Color-coded as in (a). Shaded region indicates the standard deviation over 16 randomly initialized student networks and 2 teacher networks. (c) Per-layer 2-norm of final weight updates for the different learning rules, color-coded as in (a).

4.5 Discussion

In this work, we apply the meta-learning framework to the *weight transport* problem. Here, we meta-learn both the feedback weight learning rules and the feedback pathway architecture under challenging yet tractable learning conditions. Specifically, we constrain the networks to learn entirely online, utilizing only the information available locally at each synapse.

The feedback learning rule identified in this study is a combination of the terms related to previously proposed learning rules known to improve alignment between the forward and the feedback weights (Akrout et al., 2019; Kolen and Pollack, 1994; Kunin et al., 2020). For example, Akrout et al. (2019) have shown that a Hebbian term like our f_8 coupled with a weight decay acting on a random feedback matrix $\mathbf{B}$ can effectively transpose the forward weight $\mathbf{W}$. The need for alternative, scaled normalization was raised by Kunin et al. (2020), where an Oja-like (Oja, 1982) regularization was proposed. Our term weighted by θ_{12} is similar, with key differences being (1) the normalizing term contains the layer error (not the neuron activity) and (2) we avoid any matrix multiplication, which would effectively require non-local information per synapse. Thus, our learning rule can be seen as more biologically plausible.

Finally, our term weighted by θ_1 is effectively a backpropagation term and mirrors the update done in the forward weights. Such weight update mirroring has been proposed by (Kolen and Pollack, 1994); however, here we show it also applies to *direct* feedback connections.

The flexibility of our framework allows for the incorporation of additional biological constraints, such as different activation functions or sparse feedback and forward connections, both of which have been shown to further hamper the efficacy of learning with a fixed random feedback pathway (Launay et al., 2020; Refinetti et al., 2021). Moreover, alternative formulations to standard backpropagation (BP) have shown potential for learning with initial random feedback weights (Shervani-Tabar and Rosenbaum, 2023), indicating further optimization opportunities of the learning rules operating on the forward network layers.

One significant practical limitation we encountered was the hardware (GPU) memory capacity required for our experiments. To overcome this challenge, engineering techniques such as sharding, akin to those employed in large language models, could be utilized. To support further research in this exciting area, we will publish our code as a flexible toolbox, making it accessible to the community and fostering continued exploration and development.

4.6 Appendix

4.6.1 Teacher Network Architecture and Task Complexity

Choosing the optimal size of the teacher network is crucial, especially when aiming to be in the under-parameterized regime where the student network cannot simply memorize the teacher network. However, determining this optimal size is not straightforward, as we demonstrate.

Any continuous function can be approximated arbitrarily well with a sufficiently large neural network (Funahashi, 1989). Therefore, we can consider the complexity of the teacher network in the context of the function class capacity $\mathcal{H}$ (Belkin et al., 2019). Conventional wisdom suggests that this capacity should largely be determined by the *bias-variance trade-off*, balancing data

underfitting when $\mathcal{H}$ is low, and *overfitting* when $\mathcal{H}$ is high (Geman et al., 1992; Hastie et al., 2009).

This balance can be explicitly controlled by the number of hidden units and layers in the teacher network. With a small number of teacher parameters, $\mathcal{H}$ is low, and only simple functions of the input data can be approximated, making the task too easy. Conversely, increasing the teacher network size should increase $\mathcal{H}$, thus enhancing the number of functions it can represent and increasing task complexity. We demonstrate this relationship in Figure 2.5a.

However, the relationship is not monotonic. To understand why this occurs, consider a single output unit y in the last hidden layer l of the network:

$$\bar{y}_l = \sum_{i=1}^n (W_{ji}^l \cdot a_i^{l-1}) \quad (4.24)$$

where a^{l-1} are the activations of the previous layer, and $\mathbf{W}^l$ are the weights connecting the two layers. We ignore the bias term for simplicity. Thus, the prediction is just a sample mean of all the inputs $\mathbf{a}$, weighted by $\mathbf{W}$.

Given that our input is randomly sampled from a uniform distribution, and both our weight initialization (Xavier normal (Glorot and Bengio, 2010)) and activation function ($\tanh$) are symmetric and zero-centered, the inductive bias of the network is low, and the expected value of prediction $\bar{y}_l$, i.e., $\mathbb{E}(\bar{y}_l)$, is 0. And as per the law of large numbers, $\bar{y}_l \rightarrow \mathbb{E}(\bar{y}_l)$ as $n \rightarrow \infty$.

Since in our networks there is no correlation between random variables, $\text{var}(\bar{y}_l) = \frac{\sigma^2}{n}$. We empirically show this in Figure 4.4b by measuring the variance of the y output at different network capacities. Thus, choosing too large of a teacher network can be counter-effective, as with increased network capacity, the effective expressivity of the randomly initialized teacher network collapses, allowing for a trivial task solution by an under-parameterized student network.

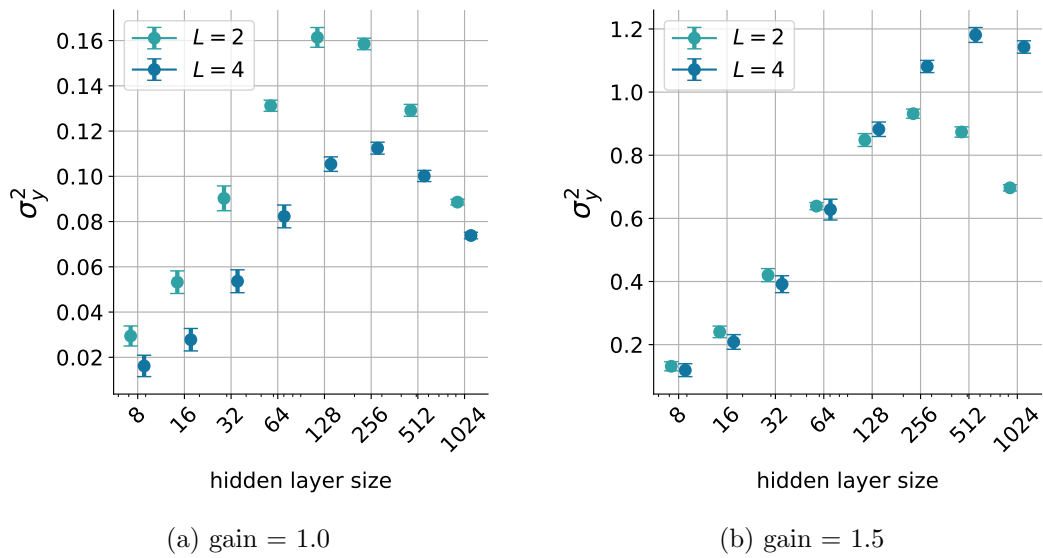


Figure 4.4: Variance of the teacher network output as a function of hidden layer size. Note the importance of the gain in weight initialization for the expressiveness of the network, as it pushes the *tanh* activation function into the non-linear regime.

Chapter 5

Conclusion

5.1 Summary

This thesis explores three interconnected open problems related to learning in the brain from various computational modelling perspectives. First, it investigates the functional implications of a novel model of inhibitory plasticity for network homeostasis, utilizing spiking neural network models inspired by the hippocampus and cortex. This modelling approach is bottom-up, where the learning rule is designed, and its emergent properties are studied. The second part of this thesis adopts a more top-down approach to address the problem of temporal credit assignment through reverse-engineering previously proposed biologically plausible mechanisms of task adaptation in pre-trained recurrent networks inspired by the motor cortex. The third and final part of this thesis takes an even more top-down modelling approach to tackle the problem of weight transport in spatial credit assignment. Here, instead of designing the learning rule or the network architecture, a set of biological constraints is defined, and both the architecture and learning rules are recovered by optimising the learning process itself. Together, these approaches offer a comprehensive exploration of learning in the brain using computational methods.

5.2 Future Work

Network homeostasis. First, on the experimental side, although we list potential implementation mechanisms, an inhibitory learning rule like the one proposed in Chapter 2 (IDIP) has yet to be experimentally verified. Thus, fully closing this experiments-theory-experiments is still necessary. Second, on the computational side, it would be of interest to study IDIP in conjunction with other homeostatic plasticity rules, as their interaction may enhance the compensatory repertoire of networks and endow them with non-trivial emergent properties. For example, weight normalization is known to accelerate training in artificial neural networks (Salimans and Kingma, 2016). What might be the effect on learning of other homeostatic strategies inspired by those we observe in the brain? Any efficiency improvements in the training stage may be especially useful under the recent trend of ever-growing, large models.

Feedback control. This thesis arrives at a time when there has been an explosion of papers linking learning in the network activity space *before* learning in the weight space, with both higher biological plausibility and more efficient learning (Alonso et al., 2022; Song et al., 2022; Innocenti et al., 2023; Meulemans et al., 2022b). The crux of these approaches has been their poor scalability due to the need for a separate relaxation phase. However, as demonstrated in Chapter 3, this relaxation phase may not be strictly necessary, opening the door for more practical applications of activity control before learning. Interestingly, the current state-of-the-art models in machine learning, transformers (Vaswani et al., 2017), attribute their generalization strides to in-context learning, where no weight changes occur at all (Von Oswald et al., 2023). This could underscore the potential for biologically inspired activity control mechanisms to further advance research efforts in both artificial and biological learning systems, leading to more effective and efficient models.

Meta learning of synaptic plasticity rules. As shown in Chapter 4 and previous work (Shervani-Tabar and Rosenbaum, 2023), gradient-based meta-learning is a powerful approach for discovering synaptic plasticity rules, but it currently scales very poorly due to the tremen-

dous compute and memory costs, as well as algorithmic instability (Choe et al., 2024). Solving these issues will be of paramount importance for the extension of our framework. Moreover, engineering solutions such as multi-GPU parallelism, currently employed to train state-of-the-art large language models, could significantly enhance the scalability of meta-learning, enabling the simultaneous discovery of both forward and backward weights in feedforward *and* recurrent biologically plausible neural networks.

5.3 Conclusion

In conclusion, this thesis provides a multifaceted exploration of learning in the brain through three distinct yet interconnected computational modelling approaches. By addressing inhibitory plasticity, temporal credit assignment, and the weight transport problem, it underscores the significance of advanced computational techniques and strong inductive biases in unraveling the mechanisms of neural learning. Looking ahead, integrating these diverse approaches will be essential for developing more effective and efficient models of learning in the brain.

Bibliography

- Abbott, L. F. and Nelson, S. B. (2000). Synaptic plasticity: taming the beast. *Nature neuroscience*, 3(11):1178–1183.
- Akrouf, M., Wilson, C., Humphreys, P., Lillicrap, T., and Tweed, D. B. (2019). Deep learning without weight transport. *Advances in neural information processing systems*, 32.
- Alemi, A., Machens, C., Deneve, S., and Slotine, J.-J. (2018). Learning nonlinear dynamics in efficient, balanced spiking networks using local plasticity rules. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32.
- Ali, A. B., Deuchars, J., Pawelzik, H., and Thomson, A. M. (1998). Ca1 pyramidal to basket and bistratified cell epsps: dual intracellular recordings in rat hippocampal slices. *The Journal of Physiology*, 507(1):201–217.
- Alle, H. and Geiger, J. R. (2006). Combined analog and action potential coding in hippocampal mossy fibers. *Science*, 311(5765):1290–1293.
- Alonso, N., Millidge, B., Krichmar, J., and Neftci, E. O. (2022). A theoretical framework for inference learning. *Advances in Neural Information Processing Systems*, 35:37335–37348.
- Andrychowicz, M., Denil, M., Gomez, S., Hoffman, M. W., Pfau, D., Schaul, T., Shillingford, B., and De Freitas, N. (2016). Learning to learn by gradient descent by gradient descent. *Advances in neural information processing systems*, 29.
- Barnes, S. J., Sammons, R. P., Jacobsen, R. I., Mackie, J., Keller, G. B., and Keck, T. (2015). Subnetwork-specific homeostatic plasticity in mouse visual cortex in vivo. *Neuron*, 86(5):1290–1303.

- Barron, H. C., Vogels, T. P., Behrens, T. E., and Ramaswami, M. (2017). Inhibitory engrams in perception and memory. *Proceedings of the National Academy of Sciences*, 114(26):6666–6674.
- Belkin, M., Hsu, D., Ma, S., and Mandal, S. (2019). Reconciling modern machine-learning practice and the classical bias–variance trade-off. *Proceedings of the National Academy of Sciences*, 116(32):15849–15854.
- Bellec, G., Scherr, F., Subramoney, A., Hajek, E., Salaj, D., Legenstein, R., and Maass, W. (2020). A solution to the learning dilemma for recurrent networks of spiking neurons. *Nature communications*, 11(1):3625.
- Bengio, S., Bengio, Y., Cloutier, J., and Gecsei, J. (2013). On the optimization of a synaptic learning rule. In *Preprints Conf. Optimality in Artificial and Biological Neural Networks*, volume 2.
- Bengio, Y., Bengio, S., and Cloutier, J. (1990). *Learning a synaptic learning rule*. Citeseer.
- Bengio, Y., Lee, D.-H., Bornschein, J., Mesnard, T., and Lin, Z. (2015). Towards biologically plausible deep learning. *arXiv preprint arXiv:1502.04156*.
- Bezaire, M. J. and Soltesz, I. (2013). Quantitative assessment of cal local circuits: knowledge base for interneuron-pyramidal cell connectivity. *Hippocampus*, 23(9):751–785.
- Bienenstock, E. L., Cooper, L. N., and Munro, P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, 2(1):32–48.
- Bliss, T. V. and Lømo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of physiology*, 232(2):331–356.
- Bock, D. D., Lee, W.-C. A., Kerlin, A. M., Andermann, M. L., Hood, G., Wetzel, A. W., Yurgenson, S., Soucy, E. R., Kim, H. S., and Reid, R. C. (2011). Network anatomy and in vivo physiology of visual cortical neurons. *Nature*, 471(7337):177–182.

- Bouhours, B., Trigo, F. F., and Marty, A. (2011). Somatic depolarization enhances gaba release in cerebellar interneurons via a calcium/protein kinase c pathway. *Journal of Neuroscience*, 31(15):5804–5815.
- Bourdoukan, R. and Deneve, S. (2015). Enforcing balance allows local supervised learning in spiking recurrent networks. *Advances in Neural Information Processing Systems*, 28.
- Briggs, F. and Usrey, W. M. (2011). Corticogeniculate feedback and visual processing in the primate. *The Journal of physiology*, 589(1):33–40.
- Buonomano, D. V. and Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annual review of neuroscience*, 21(1):149–186.
- Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*, 68(3):362–385.
- Chen, J. L., Lin, W. C., Cha, J. W., So, P. T., Kubota, Y., and Nedivi, E. (2011). Structural basis for the role of inhibition in facilitating adult brain plasticity. *Nature neuroscience*, 14(5):587.
- Chinta, L. V. and Tweed, D. B. (2012). Adaptive optimal control without weight transport. *Neural computation*, 24(6):1487–1518.
- Choe, S., Mehta, S. V., Ahn, H., Neiswanger, W., Xie, P., Strubell, E., and Xing, E. (2024). Making scalable meta learning practical. *Advances in neural information processing systems*, 36.
- Christie, J. M., Chiu, D. N., and Jahr, C. E. (2011). Ca²⁺-dependent enhancement of release by subthreshold somatic depolarization. *Nature neuroscience*, 14(1):62–68.
- Clopath, C., Vogels, T. P., Froemke, R. C., and Sprekeler, H. (2016). Receptive field formation by interacting excitatory and inhibitory synaptic plasticity. *BioRxiv*, page 066589.
- Confavreux, B., Ramesh, P., Goncalves, P. J., Macke, J. H., and Vogels, T. (2024). Meta-learning families of plasticity rules in recurrent spiking networks using simulation-based inference. *Advances in Neural Information Processing Systems*, 36.

- Confavreux, B., Zenke, F., Agnes, E., Lillicrap, T., and Vogels, T. (2020). A meta-learning approach to (re) discover plasticity rules that carve a desired function into a neural network. *Advances in Neural Information Processing Systems*, 33:16398–16408.
- Conn, A. R., Gould, N. I., and Toint, P. L. (2000). *Trust region methods*. SIAM.
- Cooke, S. F. and Bliss, T. V. (2006). Plasticity in the human central nervous system. *Brain*, 129(7):1659–1673.
- Cornford, J. H., Mercier, M. S., Leite, M., Magloire, V., Häusser, M., and Kullmann, D. M. (2019). Dendritic nmda receptors in parvalbumin neurons enable strong and stable neuronal assemblies. *Elife*, 8.
- Crick, F. (1989). The recent excitement about neural networks. *Nature*, 337(6203):129–132.
- Csicsvari, J., Czurko, A., Hirase, H., and Buzsaki, G. (1998a). Monosynaptic interactions between ca1 pyramidal cells and interneuron in the behaving rat. In *European Journal of Neuroscience*, volume 10, pages 63–63. BLACKWELL SCIENCE LTD PO BOX 88, OSNEY MEAD, OXFORD OX2 0NE, OXON, ENGLAND.
- Csicsvari, J., Hirase, H., Czurko, A., and Buzsáki, G. (1998b). Reliability and state dependence of pyramidal cell–interneuron synapses in the hippocampus: an ensemble approach in the behaving rat. *Neuron*, 21(1):179–189.
- Dale, H. (1935). Pharmacology and nerve-endings.
- Das, S., Sadanandappa, M. K., Dervan, A., Larkin, A., Lee, J. A., Sudhakaran, I. P., Priya, R., Heidari, R., Holohan, E. E., Pimentel, A., et al. (2011). Plasticity of local gabaergic interneurons drives olfactory habituation. *Proceedings of the National Academy of Sciences*, 108(36):E646–E654.
- Dauphin, Y. N., Pascanu, R., Gulcehre, C., Cho, K., Ganguli, S., and Bengio, Y. (2014). Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. *Advances in neural information processing systems*, 27.

- De Pasquale, R. and Sherman, S. M. (2011). Synaptic properties of corticocortical connections between the primary and secondary visual cortical areas in the mouse. *Journal of Neuroscience*, 31(46):16494–16506.
- Denève, S., Alemi, A., and Bourdoukan, R. (2017). The brain as an efficient and robust adaptive learner. *Neuron*, 94(5):969–977.
- Desai, N. S., Cudmore, R. H., Nelson, S. B., and Turrigiano, G. G. (2002). Critical periods for experience-dependent synaptic scaling in visual cortex. *Nature neuroscience*, 5(8):783–789.
- Desai, N. S., Rutherford, L. C., and Turrigiano, G. G. (1999). Plasticity in the intrinsic excitability of cortical pyramidal neurons. *Nature neuroscience*, 2(6):515–520.
- Deuchars, J. and Thomson, A. (1996). Ca1 pyramid-pyramid connections in rat hippocampus in vitro: dual intracellular recordings with biocytin filling. *Neuroscience*, 74(4):1009–1018.
- Dickman, D. and Wondolowski, J. (2013). Emerging links between homeostatic synaptic plasticity and neurological disease. *Frontiers in cellular neuroscience*, 7:223.
- Douglas, R. J. and Martin, K. A. (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.*, 27:419–451.
- Dupret, D., O’Neill, J., and Csicsvari, J. (2013). Dynamic reconfiguration of hippocampal interneuron circuits during spatial learning. *Neuron*, 78(1):166–180.
- D’amour, J. A. and Froemke, R. C. (2015). Inhibitory and excitatory spike-timing-dependent plasticity in the auditory cortex. *Neuron*, 86(2):514–528.
- English, D. F., McKenzie, S., Evans, T., Kim, K., Yoon, E., and Buzsáki, G. (2017). Pyramidal cell-interneuron circuit architecture and dynamics in hippocampal networks. *Neuron*, 96(2):505–520.
- Epsztein, J., Brecht, M., and Lee, A. K. (2011). Intracellular determinants of hippocampal ca1 place and silent cell activity in a novel environment. *Neuron*, 70(1):109–120.

- Felleman, D. J. and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, 1(1):1–47.
- Feulner, B., Perich, M. G., Miller, L. E., Clopath, C., and Gallego, J. A. (2022). Feedback-based motor control can guide plasticity and drive rapid learning. *bioRxiv*, pages 2022–10.
- Finn, C., Abbeel, P., and Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. In *International conference on machine learning*, pages 1126–1135. PMLR.
- Fino, E. and Yuste, R. (2011). Dense inhibitory connectivity in neocortex. *Neuron*, 69(6):1188–1203.
- Frank, C. A., Kennedy, M. J., Goold, C. P., Marek, K. W., and Davis, G. W. (2006). Mechanisms underlying the rapid induction and sustained expression of synaptic homeostasis. *Neuron*, 52(4):663–677.
- Freund, T. F. and Buzsaki, G. (1996). Interneurons of the hippocampus. *Hippocampus*, 6(4):347–470.
- Funahashi, K.-I. (1989). On the approximate realization of continuous mappings by neural networks. *Neural networks*, 2(3):183–192.
- Fyall, A. M., El-Shamayleh, Y., Choi, H., Shea-Brown, E., and Pasupathy, A. (2017). Dynamic representation of partially occluded objects in primate prefrontal and visual cortex. *Elife*, 6:e25784.
- Gainey, M. A., Aman, J. W., and Feldman, D. E. (2018). Rapid disinhibition by adjustment of pv intrinsic excitability during whisker map plasticity in mouse s1. *Journal of Neuroscience*, 38(20):4749–4761.
- Gainey, M. A. and Feldman, D. E. (2017). Multiple shared mechanisms for homeostatic plasticity in rodent somatosensory and visual cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1715):20160157.
- Gallinaro, J. V. and Rotter, S. (2018). Associative properties of structural plasticity based on firing rate homeostasis in recurrent neuronal networks. *Scientific reports*, 8(1):1–13.

- Gardner, E. and Derrida, B. (1989). Three unfinished works on the optimal storage capacity of networks. *Journal of Physics A: Mathematical and General*, 22(12):1983.
- Gasparini, S. and Magee, J. C. (2006). State-dependent dendritic computation in hippocampal cal pyramidal neurons. *Journal of Neuroscience*, 26(7):2088–2100.
- Geisler, C., Robbe, D., Zugaro, M., Sirota, A., and Buzsáki, G. (2007). Hippocampal place cell assemblies are speed-controlled oscillators. *Proceedings of the National Academy of Sciences*, 104(19):8149–8154.
- Geman, S., Bienenstock, E., and Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural computation*, 4(1):1–58.
- Gerstner, W. and Kistler, W. M. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge university press.
- Ghandour, K., Ohkawa, N., Fung, C. C. A., Asai, H., Saitoh, Y., Takekawa, T., Okubo-Suzuki, R., Soya, S., Nishizono, H., Matsuo, M., et al. (2019). Orchestrated ensemble activities constitute a hippocampal memory engram. *Nature communications*, 10(1):1–14.
- Gilbert, C. D. and Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, 14(5):350–363.
- Gilra, A. and Gerstner, W. (2017). Predicting non-linear dynamics by stable local learning in a recurrent spiking neural network. *Elife*, 6:e28295.
- Girard, P., Hupé, J., and Bullier, J. (2001). Feedforward and feedback connections between areas v1 and v2 of the monkey have similar rapid conduction velocities. *Journal of neurophysiology*, 85(3):1328–1331.
- Glazewski, S., Greenhill, S., and Fox, K. (2017). Time-course and mechanisms of homeostatic plasticity in layers 2/3 and 5 of the barrel cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1715):20160150.

- Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256. JMLR Workshop and Conference Proceedings.
- Goel, A. and Lee, H.-K. (2007). Persistence of experience-induced homeostatic synaptic plasticity through adulthood in superficial layers of mouse visual cortex. *Journal of Neuroscience*, 27(25):6692–6700.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). Convolutional networks. *Deep learning*, 2016:330–372.
- Grienberger, C., Milstein, A. D., Bittner, K. C., Romani, S., and Magee, J. C. (2017). Inhibitory suppression of heterogeneously tuned excitation enhances spatial coding in ca1 place cells. *Nature neuroscience*, 20(3):417–426.
- Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive science*, 11(1):23–63.
- Gulyás, A. I., Megias, M., Emri, Z., and Freund, T. F. (1999). Total number and ratio of excitatory and inhibitory synapses converging onto single interneurons of different types in the ca1 area of the rat hippocampus. *Journal of Neuroscience*, 19(22):10082–10097.
- Haas, J. S., Nowotny, T., and Abarbanel, H. D. (2006). Spike-timing-dependent plasticity of inhibitory synapses in the entorhinal cortex. *Journal of neurophysiology*, 96(6):3305–3313.
- Han, J.-H., Kushner, S. A., Yiu, A. P., Cole, C. J., Matynia, A., Brown, R. A., Neve, R. L., Guzowski, J. F., Silva, A. J., and Josselyn, S. A. (2007). Neuronal competition and selection during memory formation. *science*, 316(5823):457–460.
- Harris, K. D. and Shepherd, G. M. (2015). The neocortical circuit: themes and variations. *Nature neuroscience*, 18(2):170–181.
- Hartline, H. K. and Ratliff, F. (1957). Inhibitory interaction of receptor units in the eye of limulus. *The Journal of general physiology*, 40(3):357–376.

- Hartman, K. N., Pal, S. K., Burrone, J., and Murthy, V. N. (2006). Activity-dependent regulation of inhibitory synaptic transmission in hippocampal neurons. *Nature neuroscience*, 9(5):642–649.
- Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034.
- Hebb, D. O. (1962). *The organization of behavior: a neuropsychological theory*. Science Editions.
- Hengen, K. B., Lambo, M. E., Van Hooser, S. D., Katz, D. B., and Turrigiano, G. G. (2013). Firing rate homeostasis in visual cortex of freely behaving rodents. *Neuron*, 80(2):335–342.
- Hengen, K. B., Pacheco, A. T., McGregor, J. N., Van Hooser, S. D., and Turrigiano, G. G. (2016). Neuronal firing rate homeostasis is inhibited by sleep and promoted by wake. *Cell*, 165(1):180–191.
- Hennequin, G., Agnes, E. J., and Vogels, T. P. (2017). Inhibitory plasticity: balance, control, and codependence. *Annual Review of Neuroscience*, 40:557–579.
- Hinton, G., Vinyals, O., and Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*.
- Hinton, G. E. and Plaut, D. C. (1987). Using fast weights to deblur old memories. In *Proceedings of the ninth annual conference of the Cognitive Science Society*, pages 177–186.
- Hirase, H., Leinekugel, X., Czurkó, A., Csicsvari, J., and Buzsáki, G. (2001). Firing rates of hippocampal neurons are preserved during subsequent sleep episodes and modified by novel awake experience. *Proceedings of the National Academy of Sciences*, 98(16):9386–9390.
- Ho, V. M., Lee, J.-A., and Martin, K. C. (2011). The cell biology of synaptic plasticity. *Science*, 334(6056):623–628.

- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Hofer, S. B., Ko, H., Pichler, B., Vogelstein, J., Ros, H., Zeng, H., Lein, E., Lesica, N. A., and Mrsic-Flogel, T. D. (2011). Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature neuroscience*, 14(8):1045.
- Innocenti, F., Singh, R., and Buckley, C. (2023). Understanding predictive coding as a second-order trust-region method. In *ICML Workshop on Localized Learning (LLW)*.
- Jordan, J., Schmidt, M., Senn, W., and Petrovici, M. A. (2021). Evolving interpretable plasticity for spiking networks. *Elife*, 10:e66273.
- Jordan, R. and Keller, G. B. (2020). 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.
- Joseph, A. and Turrigiano, G. G. (2017). All for one but not one for all: excitatory synaptic scaling and intrinsic excitability are coregulated by camkiv, whereas inhibitory synaptic scaling is under independent control. *Journal of Neuroscience*, 37(28):6778–6785.
- Kaleb, K., Pedrosa, V., and Clopath, C. (2021). Network-centered homeostasis through inhibition maintains hippocampal spatial map and cortical circuit function. *Cell Reports*, 36(8).
- Keck, T., Scheuss, V., Jacobsen, R. I., Wierenga, C. J., Eysel, U. T., Bonhoeffer, T., and Hübener, M. (2011). Loss of sensory input causes rapid structural changes of inhibitory neurons in adult mouse visual cortex. *Neuron*, 71(5):869–882.
- Keller, A. J., Roth, M. M., and Scanziani, M. (2020). Feedback generates a second receptive field in neurons of the visual cortex. *Nature*, 582(7813):545–549.
- Kerlin, A. M., Andermann, M. L., Berezovskii, V. K., and Reid, R. C. (2010). Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron*, 67(5):858–871.
- Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

- Kirkwood, A., Rioult, M. G., and Bear, M. F. (1996). Experience-dependent modification of synaptic plasticity in visual cortex. *Nature*, 381(6582):526–528.
- Kolen, J. F. and Pollack, J. B. (1994). Backpropagation without weight transport. In *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94)*, volume 3, pages 1375–1380. IEEE.
- Korotkova, T., Fuchs, E. C., Ponomarenko, A., von Engelhardt, J., and Monyer, H. (2010). Nmda receptor ablation on parvalbumin-positive interneurons impairs hippocampal synchrony, spatial representations, and working memory. *Neuron*, 68(3):557–569.
- Krakauer, J. W., Pine, Z. M., Ghilardi, M.-F., and Ghez, C. (2000). Learning of visuomotor transformations for vectorial planning of reaching trajectories. *Journal of neuroscience*, 20(23):8916–8924.
- Kuhlman, S. J., Olivas, N. D., Tring, E., Ikrar, T., Xu, X., and Trachtenberg, J. T. (2013). A disinhibitory microcircuit initiates critical-period plasticity in the visual cortex. *Nature*, 501(7468):543–546.
- Kunin, D., Nayebi, A., Sagastuy-Brena, J., Ganguli, S., Bloom, J., and Yamins, D. (2020). Two routes to scalable credit assignment without weight symmetry. In *International Conference on Machine Learning*, pages 5511–5521. PMLR.
- Lambo, M. E. and Turrigiano, G. G. (2013). Synaptic and intrinsic homeostatic mechanisms cooperate to increase l2/3 pyramidal neuron excitability during a late phase of critical period plasticity. *Journal of Neuroscience*, 33(20):8810–8819.
- Lamprecht, R. and LeDoux, J. (2004). Structural plasticity and memory. *Nature Reviews Neuroscience*, 5(1):45–54.
- Laughlin, S. (1981). A simple coding procedure enhances a neuron’s information capacity. *Zeitschrift für Naturforschung c*, 36(9-10):910–912.
- Launay, J., Poli, I., Boniface, F., and Krzakala, F. (2020). Direct feedback alignment scales

- to modern deep learning tasks and architectures. *Advances in neural information processing systems*, 33:9346–9360.
- Le Roux, N., Cabezas, C., Böhm, U. L., and Poncer, J. C. (2013). Input-specific learning rules at excitatory synapses onto hippocampal parvalbumin-expressing interneurons. *The Journal of physiology*, 591(7):1809–1822.
- Li, L., Gainey, M. A., Goldbeck, J. E., and Feldman, D. E. (2014). Rapid homeostasis by disinhibition during whisker map plasticity. *Proceedings of the National Academy of Sciences*, 111(4):1616–1621.
- Liao, Q., Leibo, J., and Poggio, T. (2016). How important is weight symmetry in backpropagation? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 30.
- Lillicrap, T. P., Cownden, D., Tweed, D. B., and Akerman, C. J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. *Nature communications*, 7(1):13276.
- Lillicrap, T. P. and Santoro, A. (2019). Backpropagation through time and the brain. *Current opinion in neurobiology*, 55:82–89.
- Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., and Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, 21(6):335–346.
- Lindsey, J. and Litwin-Kumar, A. (2020). Learning to learn with feedback and local plasticity. *Advances in Neural Information Processing Systems*, 33:21213–21223.
- Liu, Y. H., Ghosh, A., Richards, B., Shea-Brown, E., and Lajoie, G. (2022). Beyond accuracy: generalization properties of bio-plausible temporal credit assignment rules. *Advances in Neural Information Processing Systems*, 35:23077–23097.
- Liu, Y. H., Smith, S., Mihalas, S., Shea-Brown, E., and Sümbül, U. (2021). Cell-type-specific neuromodulation guides synaptic credit assignment in a spiking neural network. *Proceedings of the National Academy of Sciences*, 118(51):e2111821118.

- Loewenstein, Y., Kuras, A., and Rumpel, S. (2011). Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. *Journal of Neuroscience*, 31(26):9481–9488.
- Losonczy, A. and Magee, J. C. (2006). Integrative properties of radial oblique dendrites in hippocampal ca1 pyramidal neurons. *Neuron*, 50(2):291–307.
- Ma, W.-p., Liu, B.-h., Li, Y.-t., Huang, Z. J., Zhang, L. I., and Tao, H. W. (2010). Visual representations by cortical somatostatin inhibitory neurons—selective but with weak and delayed responses. *Journal of Neuroscience*, 30(43):14371–14379.
- Mackwood, O. J. (2019). *Homeostatic and functional implications of interneuron plasticity*. PhD thesis, Humboldt-Universität zu Berlin, Lebenswissenschaftliche Fakultät.
- Maffei, A., Nataraj, K., Nelson, S. B., and Turrigiano, G. G. (2006). Potentiation of cortical inhibition by visual deprivation. *Nature*, 443(7107):81–84.
- Maffei, A., Nelson, S. B., and Turrigiano, G. G. (2004). Selective reconfiguration of layer 4 visual cortical circuitry by visual deprivation. *Nature neuroscience*, 7(12):1353–1359.
- Maffei, A. and Turrigiano, G. (2008). The age of plasticity: developmental regulation of synaptic plasticity in neocortical microcircuits. *Progress in brain research*, 169:211–223.
- Manita, S., Suzuki, T., Homma, C., Matsumoto, T., Odagawa, M., Yamada, K., Ota, K., Matsubara, C., Inutsuka, A., Sato, M., et al. (2015). A top-down cortical circuit for accurate sensory perception. *Neuron*, 86(5):1304–1316.
- Marschall, O., Cho, K., and Savin, C. (2020). A unified framework of online learning algorithms for training recurrent neural networks. *The Journal of Machine Learning Research*, 21(1):5320–5353.
- Martens, J. et al. (2010). Deep learning via hessian-free optimization. In *Icml*, volume 27, pages 735–742.
- Martin, S. J., Grimwood, P. D., and Morris, R. G. (2000). Synaptic plasticity and memory: an evaluation of the hypothesis. *Annual review of neuroscience*, 23(1):649–711.

- Mazzoni, P., Andersen, R. A., and Jordan, M. I. (1991). A more biologically plausible learning rule for neural networks. *Proceedings of the National Academy of Sciences*, 88(10):4433–4437.
- Metz, L., Harrison, J., Freeman, C. D., Merchant, A., Beyer, L., Bradbury, J., Agrawal, N., Poole, B., Mordatch, I., Roberts, A., et al. (2022). Velo: Training versatile learned optimizers by scaling up. *arXiv preprint arXiv:2211.09760*.
- Metz, L., Maheswaranathan, N., Nixon, J., Freeman, D., and Sohl-Dickstein, J. (2019). Understanding and correcting pathologies in the training of learned optimizers. In *International Conference on Machine Learning*, pages 4556–4565. PMLR.
- Meulemans, A., Farinha, M. T., Cervera, M. R., Sacramento, J., and Grewe, B. F. (2022a). Minimizing control for credit assignment with strong feedback. In *International Conference on Machine Learning*, pages 15458–15483. PMLR.
- Meulemans, A., Tristany Farinha, M., García Ordóñez, J., Vilimelis Aceituno, P., Sacramento, J., and Grewe, B. F. (2021). Credit assignment in neural networks through deep feedback control. *Advances in Neural Information Processing Systems*, 34:4674–4687.
- Meulemans, A., Zucchet, N., Kobayashi, S., Von Oswald, J., and Sacramento, J. (2022b). The least-control principle for local learning at equilibrium. *Advances in Neural Information Processing Systems*, 35:33603–33617.
- Meyer, H. S., Schwarz, D., Wimmer, V. C., Schmitt, A. C., Kerr, J. N., Sakmann, B., and Helmstaedter, M. (2011). Inhibitory interneurons in a cortical column form hot zones of inhibition in layers 2 and 5a. *Proceedings of the National Academy of Sciences*, 108(40):16807–16812.
- Miller, K. D. and MacKay, D. J. (1994). The role of constraints in hebbian learning. *Neural computation*, 6(1):100–126.
- Mizuseki, K. and Buzsáki, G. (2013). Preconfigured, skewed distribution of firing rates in the hippocampus and entorhinal cortex. *Cell reports*, 4(5):1010–1021.

- Mozer, M., Chauvin, Y., and Rumelhart, D. (1995). Backpropagation: Theory, architectures, and applications. *Chapter A Focused Backpropagation Algorithm for Temporal Pattern Recognition*, L. Erlbaum Associates Inc., Hillsdale, NJ, USA, pages 137–169.
- Murray, J. M. (2019). Local online learning in recurrent networks with random feedback. *Elife*, 8:e43299.
- Naumann, L. B. and Sprekeler, H. (2020). Presynaptic inhibition rapidly stabilises recurrent excitation in the face of plasticity. *bioRxiv*.
- Niell, C. M. and Stryker, M. P. (2008). Highly selective receptive fields in mouse visual cortex. *Journal of Neuroscience*, 28(30):7520–7536.
- Nøkland, A. (2016). Direct feedback alignment provides learning in deep neural networks. *Advances in neural information processing systems*, 29.
- Oja, E. (1982). Simplified neuron model as a principal component analyzer. *Journal of mathematical biology*, 15:267–273.
- O’Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Experimental neurology*, 51(1):78–109.
- O’Keefe, J. and Dostrovsky, J. (1971). The hippocampus as a spatial map: preliminary evidence from unit activity in the freely-moving rat. *Brain research*.
- O’keefe, J. and Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford: Clarendon Press.
- O’Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural computation*, 8(5):895–938.
- Pacheco, A. T., Bottorff, J., and Turrigiano, G. G. (2019). Sleep promotes downward firing rate homeostasis. *bioRxiv*, page 827832.
- Packer, A. M. and Yuste, R. (2011). Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *Journal of Neuroscience*, 31(37):13260–13271.

- Pascanu, R., Mikolov, T., and Bengio, Y. (2013). On the difficulty of training recurrent neural networks. In *International conference on machine learning*, pages 1310–1318. Pmlr.
- Payeur, A., Guerguiev, J., Zenke, F., Richards, B. A., and Naud, R. (2021). Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. *Nature neuroscience*, 24(7):1010–1019.
- Pemberton, J., Chadderton, P., and Costa, R. P. (2022). Cerebellar-driven cortical dynamics enable task acquisition, switching and consolidation. *bioRxiv*, pages 2022–11.
- Peng, Y.-R., Zeng, S.-Y., Song, H.-L., Li, M.-Y., Yamada, M. K., and Yu, X. (2010). Post-synaptic spiking homeostatically induces cell-autonomous regulation of inhibitory inputs via retrograde signaling. *Journal of Neuroscience*, 30(48):16220–16231.
- Perich, M. G., Gallego, J. A., and Miller, L. E. (2018). A neural population mechanism for rapid learning. *Neuron*, 100(4):964–976.
- Perrinet, L. U. (2010). Role of homeostasis in learning sparse representations. *Neural computation*, 22(7):1812–1836.
- Peters, A. and Payne, B. R. (1993). Numerical relationships between geniculocortical afferents and pyramidal cell modules in cat primary visual cortex. *Cerebral cortex*, 3(1):69–78.
- Peters, A., Payne, B. R., and Budd, J. (1994). A numerical analysis of the geniculocortical input to striate cortex in the monkey. *Cerebral Cortex*, 4(3):215–229.
- Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nature neuroscience*, 16(8):1068–1076.
- Pignatelli, M., Ryan, T. J., Roy, D. S., Lovett, C., Smith, L. M., Muralidhar, S., and Tonegawa, S. (2019). Engram cell excitability state determines the efficacy of memory retrieval. *Neuron*, 101(2):274–284.

- Podlaski, B. and Machens, C. K. (2020). Biological credit assignment through dynamic inversion of feedforward networks. *Advances in Neural Information Processing Systems*, 33:10065–10076.
- Poirazi, P. and Mel, B. W. (2001). Impact of active dendrites and structural plasticity on the memory capacity of neural tissue. *Neuron*, 29(3):779–796.
- Poirazi, P. and Papoutsis, A. (2020). Illuminating dendritic function with computational models. *Nature Reviews Neuroscience*, pages 1–19.
- Przybylszewski, A. W. (1998). Vision: Does top-down processing help us to see? *Current Biology*, 8(4):R135–R139.
- Ramirez, A. and Arbuckle, M. R. (2016). Synaptic plasticity: The role of learning and unlearning in addiction and beyond. *Biological psychiatry*, 80(9):e73–e75.
- Refinetti, M., d’Ascoli, S., Ohana, R., and Goldt, S. (2021). Align, then memorise: the dynamics of learning with feedback alignment. In *International Conference on Machine Learning*, pages 8925–8935. PMLR.
- Robinson, A. J. and Fallside, F. (1987). *The utility driven dynamic error propagation network*, volume 11. University of Cambridge Department of Engineering Cambridge.
- Rowan, M. J. and Christie, J. M. (2017). Rapid state-dependent alteration in kv3 channel availability drives flexible synaptic signaling dependent on somatic subthreshold depolarization. *Cell reports*, 18(8):2018–2029.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088):533–536.
- Sacramento, J., Ponte Costa, R., Bengio, Y., and Senn, W. (2018). Dendritic cortical microcircuits approximate the backpropagation algorithm. *Advances in neural information processing systems*, 31.

- Salimans, T. and Kingma, D. P. (2016). Weight normalization: A simple reparameterization to accelerate training of deep neural networks. *Advances in neural information processing systems*, 29.
- Scellier, B. and Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in computational neuroscience*, 11:24.
- Schmidhuber, J. (1987). *Evolutionary principles in self-referential learning, or on learning how to learn: the meta-meta-... hook*. PhD thesis, Technische Universität München.
- Schoenenberger, P., O'Neill, J., and Csicsvari, J. (2016). Activity-dependent plasticity of hippocampal place maps. *Nature communications*, 7:11824.
- Seung, H. S., Sompolinsky, H., and Tishby, N. (1992). Statistical mechanics of learning from examples. *Physical review A*, 45(8):6056.
- Shervani-Tabar, N. and Rosenbaum, R. (2023). Meta-learning biologically plausible plasticity rules with random feedback pathways. *Nature Communications*, 14(1):1805.
- Slomowitz, E., Styr, B., Vertkin, I., Milshtein-Parush, H., Nelken, I., Slutsky, M., and Slutsky, I. (2015). Interplay between population firing stability and single neuron dynamics in hippocampal networks. *Elife*, 4:e04378.
- Sohya, K., Kameyama, K., Yanagawa, Y., Obata, K., and Tsumoto, T. (2007). Gabaergic neurons are less selective to stimulus orientation than excitatory neurons in layer ii/iii of visual cortex, as revealed by in vivo functional ca2+ imaging in transgenic mice. *Journal of Neuroscience*, 27(8):2145–2149.
- Song, S., Sjöström, P. J., Reigl, M., Nelson, S., and Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS biology*, 3(3).
- Song, Y., Millidge, B., Salvatori, T., Lukasiewicz, T., Xu, Z., and Bogacz, R. (2022). Inferring neural activity before plasticity: A foundation for learning beyond backpropagation. *bioRxiv*, pages 2022–05.

- Steinert, J. R., Kopp-Scheinpflug, C., Baker, C., Challiss, R. J., Mistry, R., Haustein, M. D., Griffin, S. J., Tong, H., Graham, B. P., and Forsythe, I. D. (2008). Nitric oxide is a volume transmitter regulating postsynaptic excitability at a glutamatergic synapse. *Neuron*, 60(4):642–656.
- Steinert, J. R., Robinson, S. W., Tong, H., Haustein, M. D., Kopp-Scheinpflug, C., and Forsythe, I. D. (2011). Nitric oxide is an activity-dependent regulator of target neuron intrinsic excitability. *Neuron*, 71(2):291–305.
- Stemmler, M. and Koch, C. (1999). How voltage-dependent conductances can adapt to maximize the information encoded by neuronal firing rate. *Nature neuroscience*, 2(6):521–527.
- Stuart, G. J. and Spruston, N. (2015). Dendritic integration: 60 years of progress. *Nature neuroscience*, 18(12):1713–1721.
- Sweeney, Y., Kotaleski, J. H., and Hennig, M. H. (2015). A diffusive homeostatic signal maintains neural heterogeneity and responsiveness in cortical networks. *PLoS computational biology*, 11(7).
- Tamamaki, N. and Tomioka, R. (2010). Long-range gabaergic connections distributed throughout the neocortex and their possible function. *Frontiers in neuroscience*, 4:202.
- Thomson, A. M. and Radpour, S. (1991). Excitatory connections between ca1 pyramidal cells revealed by spike triggered averaging in slices of rat hippocampus are partially nmda receptor mediated. *European journal of neuroscience*, 3(6):587–601.
- Thoroughman, K. A. and Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature*, 407(6805):742–747.
- Tonegawa, S. (2019). Engram cell excitability state determines the efficacy of memory retrieval. *Neuron*, 101:1–11.
- Trouche, S., Perestenko, P. V., van de Ven, G. M., Bratley, C. T., McNamara, C. G., Campo-Urriza, N., Black, S. L., Reijmers, L. G., and Dupret, D. (2016). Recoding a cocaine-place

- memory engram to a neutral engram in the hippocampus. *Nature neuroscience*, 19(4):564–567.
- Turrigiano, G. G. (2008). The self-tuning neuron: synaptic scaling of excitatory synapses. *Cell*, 135(3):422–435.
- Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C., and Nelson, S. B. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, 391(6670):892–896.
- Turrigiano, G. G. and Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nature reviews neuroscience*, 5(2):97–107.
- Udakis, M., Pedrosa, V., Chamberlain, S. E., Clopath, C., and Mellor, J. R. (2019). Interneuron-specific plasticity at parvalbumin and somatostatin inhibitory synapses onto ca1 pyramidal neurons shapes hippocampal output. *bioRxiv*, page 774562.
- van Versendaal, D., Rajendran, R., Saiepour, M. H., Klooster, J., Smit-Rigter, L., Sommeijer, J.-P., De Zeeuw, C. I., Hofer, S. B., Heimel, J. A., and Levelt, C. N. (2012). Elimination of inhibitory synapses is a major component of adult ocular dominance plasticity. *Neuron*, 74(2):374–383.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C., and Gerstner, W. (2011). Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science*, 334(6062):1569–1573.
- Von Oswald, J., Niklasson, E., Randazzo, E., Sacramento, J., Mordvintsev, A., Zhmoginov, A., and Vladymyrov, M. (2023). Transformers learn in-context by gradient descent. In *International Conference on Machine Learning*, pages 35151–35174. PMLR.

- Werbos, P. J. (1990). Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10):1550–1560.
- Whittington, J. C. and Bogacz, R. (2017). An approximation of the error backpropagation algorithm in a predictive coding network with local hebbian synaptic plasticity. *Neural computation*, 29(5):1229–1262.
- Williams, R. J. and Zipser, D. (1989). A learning algorithm for continually running fully recurrent neural networks. *Neural computation*, 1(2):270–280.
- Wilson, D. E., Smith, G. B., Jacob, A. L., Walker, T., Dimidschstein, J., Fishell, G., and Fitzpatrick, D. (2017). Gabaergic neurons in ferret visual cortex participate in functionally specific networks. *Neuron*, 93(5):1058–1065.
- Wilson, M. A. and McNaughton, B. L. (1993). Dynamics of the hippocampal ensemble code for space. *Science*, 261(5124):1055–1058.
- Wohrer, A., Humphries, M. D., and Machens, C. K. (2013). Population-wide distributions of neural activity during perceptual decision-making. *Progress in neurobiology*, 103:156–193.
- Woodin, M. A., Ganguly, K., and Poo, M.-m. (2003). Coincident pre-and postsynaptic activity modifies gabaergic synapses by postsynaptic changes in cl- transporter activity. *Neuron*, 39(5):807–820.
- Wu, Y., Hengen, K. B., Turrigiano, G. G., and Gjorgjieva, J. (2019). Homeostatic mechanisms regulate distinct aspects of cortical circuit dynamics. *bioRxiv*, page 790410.
- Wu, Y., Ren, M., Liao, R., and Grosse, R. (2018). Understanding short-horizon bias in stochastic meta-optimization. *arXiv preprint arXiv:1803.02021*.
- Wu, Y. K., Hengen, K. B., Turrigiano, G. G., and Gjorgjieva, J. (2020). Homeostatic mechanisms regulate distinct aspects of cortical circuit dynamics. *Proceedings of the National Academy of Sciences*.

- Yassin, L., Benedetti, B. L., Jouhanneau, J.-S., Wen, J. A., Poulet, J. F., and Barth, A. L. (2010). An embedded subnetwork of highly active neurons in the neocortex. *Neuron*, 68(6):1043–1050.
- Yin, J. and Yuan, Q. (2015). Structural homeostasis in the nervous system: a balancing act for wiring plasticity and stability. *Frontiers in cellular neuroscience*, 8:439.
- Yuan, Y.-x. (2015). Recent advances in trust region algorithms. *Mathematical Programming*, 151:249–281.
- Zariwala, H. A., Madisen, L., Ahrens, K. F., Bernard, A., Lein, E. S., Jones, A. R., and Zeng, H. (2011). Visual tuning properties of genetically identified layer 2/3 neuronal types in the primary visual cortex of cre-transgenic mice. *Frontiers in systems neuroscience*, 4:162.
- Zdeborová, L. and Krzakala, F. (2016). Statistical physics of inference: Thresholds and algorithms. *Advances in Physics*, 65(5):453–552.
- Zenke, F., Hennequin, G., and Gerstner, W. (2013). Synaptic plasticity in neural networks needs homeostasis with a fast rate detector. *PLoS computational biology*, 9(11).
- Znamenskiy, P., Kim, M.-H., Muir, D. R., Iacaruso, M. F., Hofer, S. B., and Mrsic-Flogel, T. D. (2018). Functional selectivity and specific connectivity of inhibitory neurons in primary visual cortex. *bioRxiv*, page 294835.
- Zovkic, I. B. (2021). Epigenetics and memory: an expanded role for chromatin dynamics. *Current Opinion in Neurobiology*, 67:58–65.