university of groningen

faculty of science and engineering

Credit Assignment in Cortex: Backpropagation or Target Learning?

Lhea Beumer

university of groningen / faculty of science and engineering

University of Groningen

Credit Assignment in Cortex: Backpropagation or Target Learning?

Master's Thesis

To fulfill the requirements for the degree of
Master of Science in Computational Cognitive Science
at University of Groningen under the supervision of
dr. M. van Vugt (Artificial Intelligence, University of Groningen)
and
Prof. dr. B. F. Grewe (Institute of Neuroinformatics, ETH Zurich)

Lhea Beumer (s4105427)

July 5, 2025

Contents

Acknowledgments

This thesis has been the product of extensive collaboration and would not have been possible without my supervisors Pau, Sander and Benni. A huge thank you is in order for their invaluable insights, stimulating discussions, and sharp critiques. Their commitment to scientific rigor has continuously challenged me to uphold the highest academic standards. Their passion for discovery has inspired me to constantly seek deeper understanding. I'm also grateful for the incredible research group Benni has created, for how they have welcomed me with open arms and made my stay so enjoyable that I won't want to leave just yet. I also want to thank Marieke, for her great support and interest, and providing a reliable safety net if I faced difficulties. Lastly, to my loved ones, your unconditional love and unwavering support has allowed me to chase bold dreams knowing I would always return a winner, no matter the outcome. Thank you for patiently listening, genuine interests and always showing up.

Abstract

Understanding how networks of cortical neurons learn by adjusting synaptic weights remains one of the greatest frontiers of neuroscience. Machine learning theory has proposed two opposing hypotheses. Backpropagation-based (BP) algorithms that rely on minimizing error signals, and Target-based Learning (TL) algorithms in which target activity states are consolidated. This thesis provides a novel framework for experimentally distinguishing between these two theories of cortical learning and moves beyond purely theoretical arguments to concrete, data-driven validation of hierarchical learning in mouse lateral visual cortex. By showing that the reactivations of stimulus-evoked responses serve as a teaching signal and causally shape future responses, we are able to differentiate inference from learning phases. Furthermore, we reveal that the neural dynamics underlying these two phases is qualitatively distinct, which contradicts a key assumption of neuro-plausible BP approximations. Finally, results indicate that reactivation-driven plasticity is invariant to stimulus identity, leading to the hypothesis that reactivations might communicate a general teaching signal, rather than a stimulus-specific label. Collectively, we argue that these findings indicate that cortical hierarchical learning operates in a TL-compatible and BP-incompatible regime. Through the combination of causal, dynamical and representational analyses, this work furthers understanding of cortical hierarchical computation and could guide future neuro-plausible algorithm development that seek to implement the qualities of cortical computation in artificial systems.

1 Introduction

Though we have seen remarkable achievements of artificial intelligence over the past years, ability of the brain to learn from complex sensory input and improve behavior through experience still remains yet unrivaled. Nonetheless, understanding *how* networks of cortical neurons adjust their synapses to facilitate learning stands as a key open question in neuroscience. Although the mechanisms behind cellular-level synaptic plasticity are fairly well understood, they do not address how local changes align to produce coherent, goal-directed behavior across a network. Network-wide learning, such as at the level of the brain, must involve a coordination strategy that guides plasticity in a way that reflects desired behavioral outcomes.

In biology, synaptic changes are local and constrained by physiology. By contrast, ML treats synaptic updates as part of a network-level optimization problem, defined by an architecture and an explicit loss function. The backpropagation of error algorithm (BP) has emerged as the backbone for optimizing network parameters and training deep neural networks. The brain also deals with challenges in optimizing the parameters of a complex, high-dimensional network of billions of neurons and trillions of synapses. Given that the brain faces challenges with network optimization that are at the very least just as complex as artificial networks, debates have sparked over whether the brain could implement similar mechanisms for hierarchical learning in the neocortex. However, the traditional BP algorithm has faced major criticism of biological plausibility that date back as far as 1989 (Crick, 1989; Stork, 1989). In his seminal paper, Crick highlighted early concerns with bio-plausibility, noting that this algorithm assumes computational mechanisms and information flows that are in conflict with the fundamentals of brain function.

The gap between ML methods and neural circuit mechanics led to what Crick and others (Gulati et al., 2017; Stolyarova, 2018; Richards and Lillicrap, 2019; Lansdell et al., 2019) recognize as the "credit assignment" problem: how can a neuron embedded deep in cortex identify the synaptic modifications necessary to improve behavior? This problem strikes at the heart of biological learning. Unlike artificial neural networks, where BP can calculate precise synaptic weight updates for every neuron based on a global error, biological neurons have no such luxury. Biological neurons operate on local signals, ion concentrations, membrane voltages, and neurotransmitter dynamics; not on a mathematically derived precise loss function. The issue is also painfully obvious in multilayer hierarchical networks, where the consequences of a synaptic change are distributed across many downstream neurons. A single synaptic weight adjustment in a deep layer can ripple unpredictably through the system and will affect output only indirectly and often in nonlinear ways. Without a mechanism for tracing back those effects, the process by which the brain performs effective learning across multiple processing layers remains unclear. Again, we raise a fundamental question: through what learning rule does the brain assign credit where credit is due?

Over the past decades, an interdisciplinary effort spanning neuroscience and ML have put forward numerous biologically plausible learning algorithms, methods that respect the constraints of biological neural hardware while still solving credit assignment in deep networks. One strong neuro-plausible competitor to BP is Target Learning (TL). Recently coined by Aceituno et al. (2024), TL refers to a family of neuroplausible learning algorithms in which target states are consolidated. My thesis builds on the foundation laid by Aceituno et al. (2024) and examines how well TL can explain hierarchical learning in lateral visual cortex of mice compared to BP. The original analyses for the paper center around formalizing activity targets, whereas the main focus of this work is on comparing and analyzing inference and learning dynamics. The major contribution of this thesis is the implementation of using reactivations of a stimulus to predict the plasticity of that stimulus, which has been previously hinted at should be possible but never successfully implemented. Second, we analyzed whether

the dynamics of the activity state of the network during inference and learning as the two type of algorithms make different assumptions on these phenomena. Third, further analyses show that the reactivations carry a general learning signal that is invariant to stimulus identity, suggesting that they teach a functional transformation rather than a specific label. In extension to the analyses presented in Aceituno et al. (2024), we argue that cortical hierarchical learning is compatible with TL-based methods and incompatible with BP-approximations.

2 Bio-Plausible backpropagation and Target Learning Algorithms

To understand how theorists have hypothesized how the brain might solve credit assignment to implement learning, we bridge neuroscience and ML and review several biologically plausible cortical learning algorithms that can be categorized as belonging to either the BP or TL family[1], the former relying on an error to be canceled, the latter on a target activity state to be consolidated. In this chapter, we will explore the BP algorithm in detail and review the challenges associated with the neuro-biological plausibility of BP. Then, we will examine proposed neuro-plausible approximations of the BP algorithm, specifically approaches of Feedback Alignment (section 2.3.1), Dendritic Error Propagation (section 2.3.2) and Equilibrium Propagation (section 2.3.3). Afterward, we will review several learning algorithm alternatives categorized as Target Learning (TL), namely Predictive Coding (section 2.4.1), Prospective Configuration (section 2.4.2) and Deep Feedback Control (section 2.4.3). Lastly, we will highlight the difficulties in distinguishing between BP and TL when relying on theory alone and argue for the importance of experimental observation to distinguish between the two type of opposing theoretical learning algorithms.

2.1 Credit Assignments and Synaptic Weights

The strength of the connection between a pair of neurons, i.e. their synaptic weight, determines how strongly the activity of one neuron will influence the other (Markram et al., 2012). In other words, neural activity is determined by synaptic weights, and this activity determines network output. Changing synaptic weights will ultimately lead to changes in overall network behavior. During learning, the goal of the network is to improve at the task at hand and weight updates need to be coordinated such that improvement occurs. Random and unguided changes to the weights of the network will not yield meaningful behavioral gains. There must be principled ways in which a network can determine how local synaptic weight updates will contribute to the global learning objective. We previously defined this as the problem of credit assignment. A network cannot learn without solving the credit assignment problem (Bengio and Frasconi, 1993; Lillicrap et al., 2020; Aceituno et al., 2024). Effective solutions need to determine which neurons influence a task-specific network output, and be able to establish the proportional contribution of those units to the output to perform weight updates that optimize the objective (Ororbia et al., 2024). In neuroscience, the process of weight updating is referred to as synaptic plasticity.

Here we introduce the basic architecture that will serve as a common reference model for understanding how different algorithms handle credit assignment. A simple feedforward multilayer perception (MLP) (Fig. 1) can be viewed as a sequential stack of nonlinear transformations:

$$f_\Theta(x) = \{f_\ell(z^{\ell-1};\theta^\ell)\}_{\ell=1}^{L} \tag{1}$$

that is initiated by input vector $z^0 = x$. Each input transformation consists of a multiplication with the weight matrix of that layer W^ℓ and applying an element-wise activation function ϕ^ℓ to each neuron of that layer that will produce the output vector for that layer $z^\ell = f_\ell(z^{\ell-1})$. At each layer of the network, one set of transformations occurs. The pre-activation vector of a layer (h^ℓ) is obtained by multiplying the weight matrix of a layer by the post-activation vector of the previous layer: $W^\ell \cdot z^{\ell-1}$. The post-activation vector z^ℓ of the layer is then computed with the elementwise multiplication of h^ℓ with activation function ϕ^ℓ. Compute this for all layers of the network and we obtain the output $f_\Theta(x) = z^L$

[1] If the reader is interested in a review of the complete landscape of neuro-biologically plausible credit assignment algorithms, we refer to Ororbia (2023).

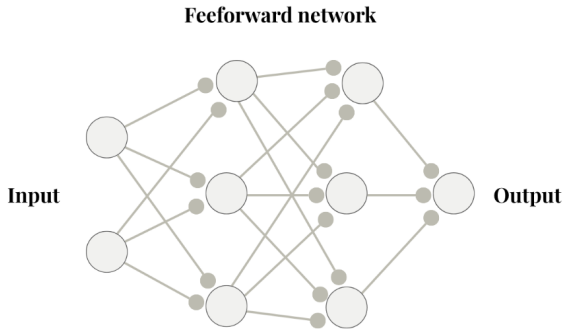


Figure 1: A simple fully-connected feedforward network (MLP) schematic that processes an input to produce an output

of the network. The objective of credit assignment is then to orchestrate updates to the network parameters $\Theta = \{W^1, W^2, \ldots, W^L\}$ to minimize a task-specific loss function. This loss function is a metric to determine the quality of the output w.r.t. the task. Both the loss and activation functions are selected by the system designer on the basis of their suitability for the specific task the network is intended to perform.

2.2 Backpropagation

In BP, to update the synaptic weights in the network, gradients are derived from the loss function $\mathcal{L}(y, z^L)$ given the output z^L formed by each weight matrix W^ℓ that are in the set of our parameters Θ and the desired output y. For any given layer (ℓ) in our network, we recursively apply the chain rule to compute these gradients $\frac{\partial \mathcal{L}}{\partial W^\ell}$. Each gradient quantifies how sensitive the loss is to a change in each weight, thereby indicating which weights should be increased or decreased and by how much to improve performance. The signal begins at the output and travels backward through the network (hence the name 'backpropagation'). Gradients are thus computed layer by layer using the structure of the computational graph of the network. For the final layer, the weight update ΔW^L is computed with:

$$\Delta W^L = \frac{\partial \mathcal{L}(y, z^L)}{\partial h^L} (z^{L-1})^\top \tag{2}$$

which can be broken down into an error signal δ^ℓ that quantifies the quality of the output given the loss function and the elementwise product between this error signal and the activation derivative $\phi'(h^L)$ that says how much each neuron output changes when its input changes:

$$\delta = \frac{\partial \mathcal{L}}{\partial h^L} = \frac{\partial \mathcal{L}}{\partial z^L} \odot \phi'(h^L) \tag{3}$$

where $\phi'(h^L)$ denotes the derivative of the activation function evaluated at the pre-activation vector h^L of layer L. Therefore, the weight update becomes:

$$\Delta W^L = \delta^L (z^{L-1})^\top. \tag{4}$$

For earlier layers $\ell < L$, the error signals δ^ℓ are computed similarly with:

$$\delta^\ell = (W^{\ell+1})^\top \delta^{\ell+1} \odot \phi'(h^\ell) \tag{5}$$

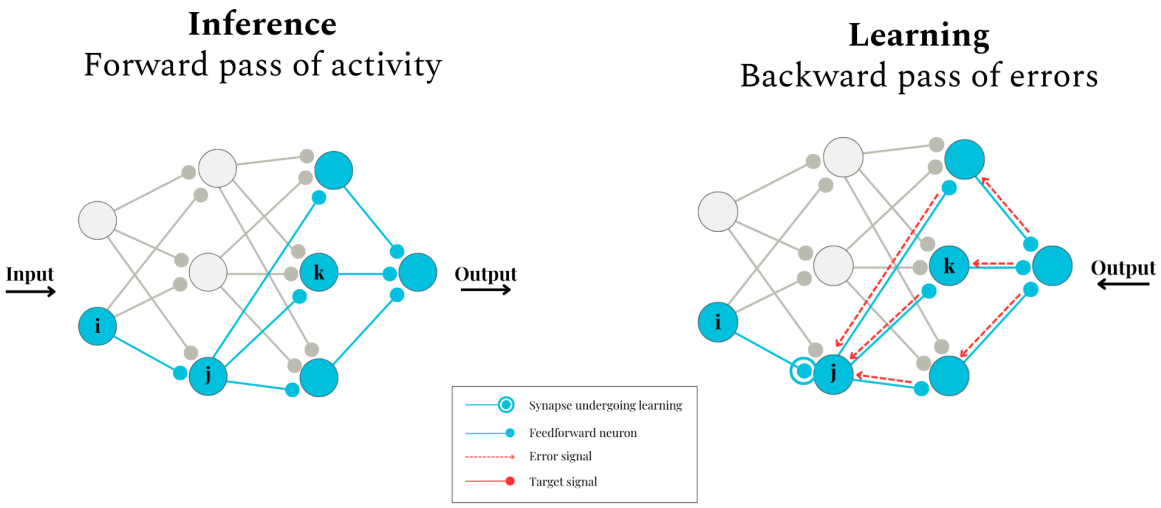


Figure 2: Visualization of the backpropagation learning algorithm. **Left**: schematic of the feedforward inference phase in which the network produces a blue activity state. **Right**: schematic of feedback learning phase in which the network computes small errors given the output, which are propagated backward through the network to each synapse while being in the same blue activity state.

The purpose of the term $\delta^{\ell+1} \odot \phi'(h^{\ell})$ is to scale the error with the responsiveness of each neuron. Neurons whose outputs are near the saturation regions of their activation functions (close to the upper or lower limit) will have small gradients because the derivative of the activation function is near zero in those regions. The corresponding weight update for this layer, as we have seen in equation 4, then becomes:

$$\Delta W^{\ell} = \delta^{\ell}(z^{\ell-1})^{\top} \tag{6}$$

Figure 2 shows a visualization of how BP works. We have separated the inference (feedforward) phase from the learning (feedback) phase. Described by the formalizations above, the network first processes the input during inference, after which the gradients are computed and propagated backwards through the network to communicate to a synaptic weight the direction and magnitude of the necessary update. Crucially, the activity is in the same, or very similar, state across inference and learning as the error signal that is backpropagated is relatively small (Bengio et al., 1994; Meulemans et al., 2022).

So, backpropagation is the recursive application of the chain rule that enables gradient-based optimization in multilayer networks. But at a deeper level, the algorithm has powerful implications about how error signals are distributed, how representation learning occurs and how parameter updates emerge from functional dependencies between layers (Rumelhart et al., 1986). BP traces the causal influence of each parameter back through the computational graph of the network. By computing layer-by-layer gradients, BP ensures that every parameter is updated proportionally to their individual (indirect) contribution to the overall loss. Each parameter is updated based on the extent to which a change would influence future outputs, given the same input. The combination of mathematical rigor and computational efficiency has placed the BP learning algorithm at the center stage of all of the extraordinary performances of modern artificial intelligence.

2.2.1 Problems with BP

As alluded to earlier, backpropagation faces serious critiques regarding biological plausibility due to reliance on computational mechanisms and information flows that seem unlikely to be implemented by the brain. Below, we review the key criteria that are incompatible with the current understanding of brain function.

Global Feedback & Vanishing Gradients: Early on, researchers realized that BP demands a global error signal that flows backward through the network. BP assumes that a single feedback signal can be distributed across multiple layers. In theory, this provides a precise credit assignment for all neurons given their contribution to the overall error. However, implementing this type of global error broadcast in a biological network seems far from trivial. Every single neuron in the hierarchy would need information about the difference between a desirable outcome ant the actual output of the network. Shortly after the conception in 1986 (Rumelhart and McClelland, 1986), critics like Crick (1989) and Stork (1989) argued that the brain was unlikely to implement this kind of precise but centralized error broadcast.

Moreover, as networks grew deeper in the late eighties and nineties, researchers encountered problems with vanishing or exploding gradients (Bengio and Frasconi, 1993). The network is prone to instability and gradients can explode or vanish as they propagate through many layers since the feedback chain is long. Intuitively, each layer only passes on a fraction of the original error when gradients are multiplied with local derivations of the activations and weights of that layer. By the time the error signal will have reached the first layers, much of the useful learning signal has faded away (Hochreiter et al., 2001). Even if the brain could send a feedback message from output neurons back to neurons many synapses away, that message might deplete or get overwhelmed by noise. Whether the brain could maintain a single precise error signal across many stages of processing remains unclear. Indeed, the brain appears to be more robust and not suffer from the exact gradient issues that purely global signals do. The brain might instead rely on more local information or modulatory signals that provide only coarse feedback and not detailed layer-by-layer error corrections (Crick, 1989; Ororbia et al., 2024).

Thus, BP computes synaptic weight changes using information that is not locally accessible by the neuron. This reliance on global error signal that is dependent on the activity of many neurons is directly at odds with the constraints of the brain, where each neuron relies solely on information received from directly connected neurons. Biological neurons rely on processing information that is local to them in both space and time. Taken together, these critiques underscore how the need for a global error broadcast is hard to reconcile with the distributed, decentralized nature of brain information processing.

Weight Transport Problem: Another fundamental bio-plausible limitation is that BP requires the exact same weight matrix for the backward pass as used for the forward pass (Grossberg, 1987). The backward pass multiplies errors with the transpose of the forward weights. In simulation this is trivial, we simply copy the weight values. However, in a biological network, the synaptic connections are typically non-symmetric and there is no obvious mechanism for a neuron to know the exact strength of a synapse projecting in the opposite direction. A weight W_{ij} from neuron i to j (Fig. 2) is updated in proportion to weight W_{jk} from j to some higher neuron k (as part of the error gradient). This implies neuron i must somehow use the value of W_{jk} in its own weight update. Therefore, information about W_{jk} is required to be accessible at synapse W_{ij} and this must be true for all neurons embedded in the network. Any neuron needs to know information about all of the downstream weights of all the other neurons in the network. In other words, the downstream weights must be transported in order to make an optimal update to the weight of neuron i (Lillicrap et al., 2016).

In biological networks, neuron i cannot access the synaptic strength between neurons j and k, as that connection is located on the dendrite of a different neuron. Each synapse has access only to the strength of its own connection (Akrout et al., 2019). Moreover, the dedicated feedforward and feedback pathways present in the brain are not the exact mirror of each other. The weight in direction $i-j$ does not need to be the exact same as the weight in direction $j-i$, and realistically, given the messiness inherent to biology (as opposed to mathematics), weight are likely not exactly mirrored. Even more so, they are thought to serve different functional roles. While feedforward weights ($i-j$) simply process network input, the feedback weights ($j-i$) are often associated to more complicated cognitive processes such as attention, prediction or context signals (Zagha, 2020; Semedo et al., 2022). Moreover, these feedback connections are also able to influence the neural activities produced in the forward pass directly (Lillicrap et al., 2020). This is unlike BP, in which feedback will only alter the weights, not the activity states of neurons.

Inference - Learning Dependence In BP, the synaptic weight update for a neuron depends on the activation of that neuron from the forward pass. This means that the learning and inference phase are tightly coupled in a two-phase computation. Learning in BP cannot proceed until an inference step has been finished, and afterward uses the outcome of this inference to learn. This means that learning is entirely dependent on the prior inference and cannot occur simultaneously with inference. Weights are only updated after the network has fully processed the input and produced an output. Moreover, BP also uses the activations of neurons during the forward pass to compute weight updates. This dependency implies that the network needs a mechanism to store all intermediate activations until the backward phase occurs.

Thus, BP operates in distinct phases; forward, then backward. This distinction of phases of naturally leans into another implementation issue that limit parallel or online processing as a consequence of the feedforward structure. In a standard deep network, layer ℓ cannot process new input until layer $\ell-1$ has produced an output. Furthermore, layer ℓ can also not perform weight updates until all layers that come before have finished updating their weights, thus having to wait for the backpropagated error to reach layer ℓ. By contrast, biological neurons operate continuously and in parallel.

2.3 Bio-plausible BP

With these limitations, widespread consensus holds that the brain cannot feasible implement backpropagation as currently implemented in artificial neural networks, since doing so would violate core biological constraints. However, BP-like solutions have emerged that approximate gradient descent in a way that is more consistent with the biological principles of the brain. Below, we will review several of such gradient-based learning algorithms, Feedback Alignment, Dendritic Error Propagation and Equilibrium Propagation, that are derived from BP and yet are more compatible with brain function.

2.3.1 Feedback Alignment

One of the earliest breakthroughs in making BP more neuro-plausible was the discovery that learning can occur even when feedback weights are random and fixed. Lillicrap et al. (2016) showed that by relaxing the strict weight symmetry requirement of BP, networks can still learn effectively. When errors are propagated backward through random feedback synapses (direction W_{ji}, Fig. 2) the network gradually learns to align the feedforward pathways (direction W_{ij}) in such a way that supports effective credit assignment. This finding reduced concerns with the bio-implausibility of the weight transport problem and suggests that cortical feedback connections could be different from feedforward ones

without negatively influencing learning. The update rule is similar to that of equation 4, but introduces a fixed, randomly initialized feedback matrix B_ℓ. These parameters are non-trainable and thus remain constant over time. Weight updates for a given layer ℓ are computed using

$$\Delta W^\ell = \hat{\delta}^\ell (z^{l-1})^\top \tag{7}$$

Crucially, the error signal for each layer $\hat{\delta}^\ell$ is shaped by the structure of the corresponding feedback matrix B^ℓ. As a result, error signals are no longer constrained by symmetric feedback and take on different values from those computed with backpropagation.
The error signals are recursively defined as

$$\hat{\delta}^\ell = B^\ell \frac{\partial L}{\partial h^{\ell+1}} = (B^\ell \hat{\delta}^{\ell+1}) \odot \phi'(h^\ell) \tag{8}$$

This recursive form replaces the transpose of the forward weight matrix with B^ℓ, preserving the general structure of error propagation while relaxing the biological implausibility of symmetric weight transport. Substituting this into equation 7 yields the final weight update:

$$\Delta W^\ell = \left(B^\ell \hat{\delta}^{\ell+1} \odot \phi'(h^\ell) \right) (z^{\ell-1})^\top \tag{9}$$

Feedback Alignment, however, does not offer a complete solution to the credit assignment problem. While it relaxes the biologically implausible requirement for symmetric backward weights, the generated learning signals remain relatively global, and performance with purely random feedback matrices often lacks robustness across different architectures and tasks (Bartunov et al., 2018; Moskovitz et al., 2018; Launay et al., 2019a; Crafton et al., 2019). To overcome these limitations, subsequent research has explored mechanisms for improving feedback signals. One notable extension is Direct Feedback Alignment (DFA) (Nøkland, 2016), which transmits error signals from the output layer directly to each hidden layer, bypassing intermediate layers entirely. This strategy enables layer-wise parallel learning and mitigates the vanishing and exploding gradient problems discussed before.

$$\Delta W^\ell = \left(B^\ell \hat{\delta}^L \odot \phi'(h^\ell) \right) (z^{\ell-1})^\top \tag{10}$$

Here, B^ℓ is a fixed random feedback matrix that maps output-layer error signals $\hat{\delta}^L$ directly to hidden layer ℓ and $\phi'(h^\ell)$ denotes the derivative of the activation function at that layer. This formulation ensures that each hidden layer receives an error signal without relying on recursive error propagation through preceding layers. Nevertheless, the applicability of DFA is limited in other architectures such as CNNs where effectively aligning the random feedback weights for learning is more challenging (Bartunov et al., 2018; Launay et al., 2019b).
The conceptual difference between networks like perceptrons that we have discussed so far (i.e. Fig. 1) and cortical networks made up of pyramidal neurons (PNs) further highlight the abstraction gap in models such as DFA. Perceptrons, as used in artificial neural networks, aggregate input, apply a nonlinearity, and propagate signals forward in a uniform and mathematically well-defined manner. Error signals, in turn, are backpropagated using idealized gradients computed across layers. In contrast, PNs in the neocortex exhibit complex compartmentalized processing, receiving bottom-up input at basal dendrites and top-down feedback signals at apical dendrites (Larkum, 2013). These structurally distinct compartments suggest a form of segregated signal flow that is not easily captured by simple feedforward models. The direct transmission of error signals to hidden layers in DFA resembles the top-down signal transmission onto apical dendrites. However, unlike perceptrons, PNs integrate feedback and feedforward signals in a nonlinear, state-dependent manner that supports more sophisticated

credit assignment mechanisms than those modeled by DFA (Sacramento et al., 2018; Meulemans et al., 2022). This fundamental difference explains why biologically plausible learning frameworks often seek to incorporate dendritic processing and compartment-specific plasticity, rather than relying on the uniform, architecture-agnostic feedback strategies that we have discussed so far.

2.3.2 Dendritic Error Propagation

Sacramento et al. (2018) propose a different method for bio-plausible credit assignment through an approximation of backpropagation. This work draws inspiration from the architecture of the cortical microcircuit by mapping computational roles to dendritic compartments of two neuron types found in the neocortex. The PNs are compartmentalized into basal, apical and somatic regions, representing the processing of feedforward and feedback signals and producing a neural firing rate, respectively. Second, a population of interneurons is introduced that cancel top-down inputs by isolating the error signals at the apical dendrites of PNs. Modeled as two-compartment units with dendritic and somatic part, these interneurons are primarily driven by PNs from the same layer and project back to the apical dendrites of those same PNs.
Synaptic plasticity is formalized with the general dendritic predictive plasticity rule:

$$\Delta W = \eta(\phi(u) - \phi(v))r \tag{11}$$

that contains a presynaptic input to the neuron r, a learning rate of η and an activation function ϕ on each of the distinct compartment activities u and v. Importantly, each of the different connections between compartments of the dendritic error prop method follows their own weight update rule.
The feedforward basal weights between two pyramidal neurons (PP) between layer $l-1$ and l are updated as follows:

$$\Delta W_{l,l-1}^{PP} = \eta_{l,l-1}^{PP}\left(\phi(u_l^P) - \phi(\hat{v}_{B,l}^P)\right)\left(r_{l-1}^P\right)^T \tag{12}$$

Here, the basal dendrite $\hat{v}_{B,l}^P$ predicts the somatic potential u_l^P and the synaptic weight is adjusted to minimize this difference. Second, the connections between interneurons (I) and pyramidal (P) neurons are updated with

$$\Delta W_{l,l}^{IP} = \eta_{l,l}^{IP}\left(\phi(u_l^I) - \phi(\hat{v}_l^I)\right)\left(r_l^P\right)^T \tag{13}$$

in which u_l^I represents the interneuron somatic potential, $\hat{v}_l^I$ the predicted somatic effect of the dendritic input and r_l^P the presynaptic firing-rate of a PN. This rule trains interneurons to predict the activity of the next higher pyramidal layer and adjusts the inputs from pyramidal cells so that the interneurons replicate higher layer activity. Finally, connections between the apical dendrites of PNs of layer ℓ and the interneurons are updated through

$$\Delta W_{l,l}^{PI} = \eta_{l,l}^{PI}\left(v_{\text{rest}} - v_{A,l}^P\right)\left(r_l^I\right)^T \tag{14}$$

in which v_{rest} represents the desired apical activity in the absence of top-down input, $v_{A,l}^P$ the activity in the apical compartment of PNs arising from a combination of top-down feedback and lateral interneuron input and r_l^I denoting the interneuron activity. This rule ensures that interneurons cancel the predictable portion of top-down input, effectively zeroing out apical dendritic voltage in the absence of an external teaching signal.
In contrast to the learning algorithms previously discussed, this method takes direct inspiration from the structure of real neurons and works to implement these insights to make the algorithms more neuro-plausible. However, aside from taking inspiration of the structure of neurons, other approaches

aim to model the dynamical nature of the brain where neural states evolve continuously over time in response to inputs.

2.3.3 Equilibrium Propagation

Finally, we will discuss the last of the backpropagation-based neuro-plausible approximations, before moving on to review Target Learning. Equilibrium propagation (EqProp) models the brain as a dynamical system that naturally evolves toward energy minima, replicating the energy-minimizing dynamics observed in biological circuits (Scellier and Bengio, 2016). Learning emerges from tiny perturbations in the dynamics caused by a small influence from a feedback signal, without explicitly requiring backward passes. Crucially, synaptic weight changes depend only on locally available information and the weight updates require only one type of computation.
EqProp operates on energy-based models, in particular continuous Hopfield networks, where neural states u evolve to minimize energy function $E(u,W)$ with W representing synaptic weights. The learning procedure happens in two phases. In the free phase ($\beta = 0$) the network is clamped only to the input and allowed to relax to a free steady state u^0, minimizing the internal energy function E. Then, a small perturbation ($\beta > 0$) is applied to the output neurons to encourage movement toward a target value y, after which the network relaxes again to reach a new perturbed steady state u^β.
The weight update rule is given by:

$$\Delta W = -\eta \frac{1}{\beta} \left(\frac{\partial E}{\partial W}(u^\beta) - \frac{\partial E}{\partial W}(u^0) \right) \tag{15}$$

This rule implies that the synaptic update is proportional to the difference in the energy gradients with respect to the weights before and after the target-driven perturbation.
For a Hopfield-type energy function E with the general form of:

$$E(u,\theta) = \frac{1}{2}\sum_i u_i^2 - \frac{1}{2}\sum_{i\neq j} W_{ij}\phi(u_i)\phi(u_j) - \sum_i b_i \phi(u_i) \tag{16}$$

with activation function ϕ the weight update becomes:

$$\Delta W_{ij} = \frac{1}{\beta}\left(\phi(u_i^\beta)\phi(u_j^\beta) - \phi(u_i^0)\phi(u_j^0)\right) \tag{17}$$

where u^0 represents the neural state of the network at equilibrium in the free phase, u^β represents the neural state in the weakly clamped (nudged) phase, ϕ is an activation function, β a small scalar controlling the strength of output nudging and η is the learning rate. The function has a local Hebbian form in that synapses are updated proportionally to the difference in co-activation between the nudged and free states.

2.4 Target Learning

So far we have explained the BP algorithm and discussed how BP can be approximated to improve neuro-plausibility. Importantly, in all of the discussed algorithms, the network dynamics during learning and inference operate in the same activity regime (Fig. 2). However, these approximations have highlighted merely one approach of learning algorithms. The term "Target Learning" was coined recently by Aceituno et al. (2024) and refers to a different family of algorithms that is functionally distinct from backpropagation and its bio-plausible derivations. This alternative class relies on imposing target activities on neurons, rather than communicating explicit errors defined by a loss function.

During learning, the network assumes a distinct target activity state that is communicated with feedback or control signals (Fig. 3). Weight updates are then performed so that this target activity is consolidated. Importantly, synaptic plasticity in this family of algorithms explicitly relies only on the information that is accessible by the neuron, i.e. the activity of other directly connected neurons. This framework is mathematically distinct from BP as TL requires the network to operate in distinct activity phases, one in which normal inference happens and one in which target activity states are present.

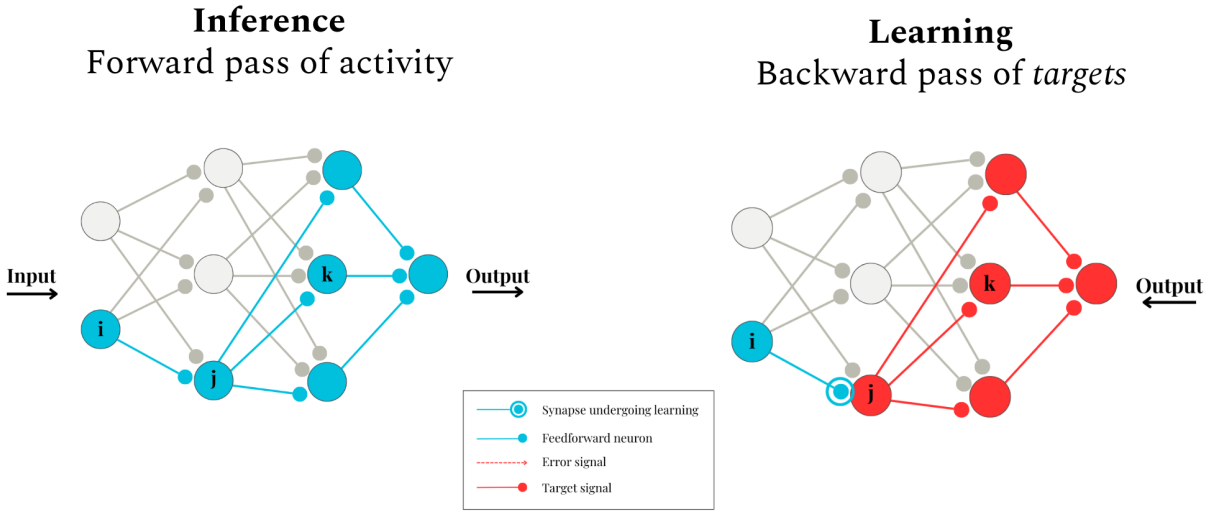


Figure 3: Visualization of Target Learning algorithms. **Left**: schematic of the feedforward inference phase in which the network produces a blue activity state. **Right**: schematic of feedback learning phase in which the network adopts a red target activity state, so that each synapse adapts in order to consolidate the activity state.

Also, TL-based methods stand in contrast to BP as they leverage second-order optimization principles. In optimization theory, first-order methods like gradient descent use only the local slope of the loss surface to adjust parameters (Nocedal and Wright, 2006). Second-order methods estimate curvature information of the loss landscape to inform weight updates, such as the Hessian or Gauss-Newton matrix (Sacramento et al., 2018). Second-order methods are favorable as they typically accelerate convergence and yield more robust weight updates. However, such methods come with increased computational cost due to the incorporation of richer information about the loss landscape, but enables more informed and often larger updates during training (Nocedal and Wright, 2006; Bottou et al., 2018).

In what follows, we will present three different target-based learning algorithms that fall under the TL umbrella: Predictive Coding, Prospective Configuration and Deep-Feedback Control (DFC), each taking on new perspectives on how learning can occur in a network that relies on top-down target activities communicated through feedback, while respecting the constraints of biology. We shall then explore the challenges between making a full distinction between BP- and TL-based algorithms when relying on theory alone, thereby arguing for the need of experimental observations to establish which of the two families is actually a better model for explaining how the cortex implements deep learning algorithms.

2.4.1 Predictive Coding

To start, Rao and Ballard (1999) propose a hierarchical predictive coding framework in which each cortical layer predicts activity in the layer below via feedback connections, while transmitting residual prediction errors upward via feedforward pathways. This model treats the brain as a probabilistic generative system, continuously comparing internal predictions with incoming sensory data to update its internal model, consistent with the Bayesian brain hypothesis (Friston, 2012; Rahnev, 2019). Predictive coding models two distinct neuron types: error neurons, which encode mismatches between predicted and observed activity, and state neurons, which represent inferred causes or latent variables. Prediction errors are propagated hierarchically and drive synaptic updates that minimize discrepancies across layers. Empirical evidence supports this mechanism, showing neural responses in early sensory areas, such as V1, reflect differences between expected and actual sensory input during perceptual tasks (Wacongne et al., 2011; Spratling, 2019).
Formally, weight updates in the network take a different form than discussed in the backpropagation section. The predictive model generates an input I from latent causes r via weights W and an activation function f and some noise n:

$$I = f(Wr) + n \tag{18}$$

Given the explicit hierarchy of the model, we iteratively generate top-down predictions r^{td} on the input from lower layers with a similar function:

$$r = r^{td} + n^{td} = f(W^{\ell} r^{\ell}) + n^{td} \tag{19}$$

in which $f(W^{\ell} r^{\ell})$ represents the top-down input prediction of another layer ℓ given that r^{ℓ} itself represents latent causes of the higher layers. We want to optimize the parameters of our weights W at each layer ℓ. Under assumption of Gaussian noise σ, the negative log-likelihood is the sum of weighted squared errors

$$E_1 = \frac{1}{\sigma^2} \left\| I - f(Wr) \right\|^2 + \frac{1}{\sigma_{\text{td}}^2} \left\| r - f(W^{\ell} r^{\ell}) \right\|^2. \tag{20}$$

This optimization function consists of several important terms. The first, $\frac{1}{\sigma^2}(I - f(Wr))^T$, quantifies how well the internal representation r and weight matrix W reconstruct the input I. Minimizing this term pushes the model toward accurate reconstructions. The second term, $\frac{1}{\sigma_{td}^2} r - f(W^{\ell} r^{\ell})$, quantifies how closely the representation matches predictions from higher layers to enforce consistency across layers. Minimizing this ensures hierarchical coherence in representations. Thus, each term measures prediction error in the feedforward and feedback pathways, weighted by the inverse noise variances. Then we can write out the full energy function E

$$E = E_1 + g(r) + h(W) \tag{21}$$

for which $g(r)$ and $h(W)$ terms represent prior beliefs or constraints about the representation and weights and enforces desirable properties such as sparsity for r or weight decay for W.
By performing gradient descent on E w.r.t. the weights W, Rao and Ballard (1999) derive a local Hebbian-style update rule

$$\Delta W = \frac{1}{\sigma^2}(I - f(Wr)) r^T - \lambda W \tag{22}$$

where $(I - f(Wr))$ is the local prediction error (postsynaptic signal) and r is the presynaptic activity. The term λW represents a Gaussian weight-decay prior.
In summary, in a predictive coding network, each synapse of W_{ij} is adjusted in proportion to the product of the error signal at its postsynaptic neuron and the presynaptic activity, minus a decay term. Importantly, this update depends only on locally available information. No global error signals or separate error-propagation pathways are required beyond the immediate prediction-error feedback at each layer.

2.4.2 Prospective Configuration

For the second TL algorithm, we discuss how Song et al. (2024) propose a new learning principle termed prospective configuration in which the network first infers a target activity pattern that should result from learning, after which synaptic plasticity consolidates these activity patterns. The authors argue that prospective configuration aligns better with neural dynamics observed in biological systems. Moreover, they argue that this method provides superior learning efficiency and robustness against interference.
The essence of prospective configuration is explained within energy-based networks, particularly the predictive coding models that we have previously discussed. These type of networks also minimize an energy function E which quantifies prediction errors within the network. This energy function E of a given layer ℓ is of the form

$$E^l = \frac{1}{2}(u^\ell - W^{\ell-1} f(u^{\ell-1}))^2 \tag{23}$$

in which the term $W^{\ell-1}$ denotes the weights of the previous layer and u^ℓ and $u^{\ell-1}$ denote the activities of a layer and the previous layer, respectively.
Minimizing this energy function happens in two distinct stages. First, during relaxation, neurons adjust their activities to minimize a prediction error by moving in the gradient direction

$$\Delta u = -\eta \frac{\partial E}{\partial u^\ell} \tag{24}$$

thereby giving a description of the neural dynamics of the network with a learning rate of η. After the neural activities have converged to their prospective states, synaptic weights are updated to consolidate this prospective configuration with

$$\Delta W = -\alpha \frac{\partial E}{\partial W}. \tag{25}$$

Thus, this prospective configuration principle is distinguished by its two-step learning rule: first optimizing neural activities based on local errors, then updating synaptic weights to stabilize this optimized configuration.

2.4.3 Deep Feedback Control

The last TL-based algorithm we will discuss is the deep feedback control (DFC) method (Meulemans et al., 2021). Inspired by control theory, DFC employs dynamic controllers that drive neural activity toward target outputs. Synaptic plasticity arises from local activity-dependent changes, and learning occurs in two distinct phases (similar to the previously discussed prospective configuration method). However, DFC differs by explicitly incorporating a feedback controller. Here, weight updates also serve to consolidate target activity patterns shaped by the controller.

The DFC method also takes direct inspiration from the structure of cortical PNs, and models distinct feedforward (basal, W) and feedback (apical, Q) weights. The DFC operates on a continuous, dynamical neural network model for which a neuron's pre-activations v_i at time t with time-constant τ_v is described by the neuron dynamics $\tau_v \frac{d}{dt} v_i(t)$. The dynamics of the network activity $\Delta u^\ell(t)$ at time t and layer l that integrates feedforward (W) and feedback (Q) inputs are given by

$$\Delta u^\ell(t) = -u^\ell(t) + W^\ell \phi(u^{\ell-1}(t)) + Q^\ell u(t) \tag{26}$$

in which the feedback signal $u(t)$ is multiplied with the feedback weights of that layer Q^ℓ. During learning, the feedback controller actively drives the the output of the network r_L toward a target output r_L^* with learning rate η as follows:

$$r_L^* = r_L^- - \eta \frac{\partial L(r_L, y)}{\partial (r_L)}\bigg|_{r_L = r_L^-}. \tag{27}$$

Synaptic weight updates are computed using local activity-based rules. For clarity and to keep the review for this thesis focused, we will discuss only the computation of the forward weights W^ℓ, which are defined as follows:

$$\Delta W^\ell = (\phi(u^\ell(t)) - \phi(W^\ell r^{\ell-1}(t)) r^{\ell-1}(t))^T. \tag{28}$$

So, the update depends locally on the difference between neuron activations driven by feedback and activations predicted solely from feedforward inputs, multiplied by presynaptic activity $r^{l-1}(t)$. Interestingly, this method relies on second-order Gaussian-Newton optimization, rather than the first-order method that BP uses (Sacramento et al., 2018).

2.5 Convergence of Learning Paradigms

Although BP-based and TL-based algorithms originate from different theoretical foundations, both often produce similar computational outcomes and circuit-level implementations. This convergence complicates efforts to distinguish between the two approaches using theoretical analysis alone, leaving open the possibility that biological systems may implement BP, TL, or a hybrid approach (Aceituno et al., 2024).

Recent studies have found that BP and TL are not mutually exclusive and can emerge from shared mathematical frameworks. Equilibrium Propagation (Scellier and Bengio, 2016) and Predictive Coding (Whittington and Bogacz, 2017) frame learning as a dynamic inference process where neural activity settles to a fixed point. In these models, gradients are approximated by comparing two steady states, a free state and a nudged state. This comparison procedure resembles TL terminology but mathematically approximates BP. Moreover, Feedback Alignment (Lillicrap et al., 2016) and Deep Feedback Control (Meulemans et al., 2021) demonstrate that error signals do not need to be perfectly backpropagated to achieve learning. Approximate error directions, delivered by biologically plausible feedback circuits, suffice for successful learning even under local plasticity rules that resemble TL. Moreover, prospective configuration (Song et al., 2024) explicitly bridges the two paradigms. In this framework, networks first infer neural activities that would reduce output error (TL-like), and then update weights to consolidate those states, which functionally resembles BP, thereby blurring the boundary between target consolidation and error propagation. Moreover, in the limit, when the network is in a near-optimal parameter state, the weight updates for BP and TL are the same. The differences between actual and target activity become small and therefore align with small BP-computed

gradients. Thus, in the late stages of learning, the observable distinction between BP and TL effectively disappears.
In addition to mathematical convergence, both BP and TL can be instantiated within similar circuit architectures. A powerful example is provided by Sacramento et al. (2018), who describe that a cortical microcircuit model can approximate both predictive coding (TL) and error-backpropagation. Moreover, both architectures include elements of biological neural circuits such as recurrent dynamics, feedback connections and inhibitory interneurons, indicating that none are exclusive to any single algorithmic interpretation. The same circuit motifs can support BP-like gradient computation or TL-like target inference, depending on how they are read out and how plasticity is implemented. Thus, this architectural flexibility makes that observing a particular circuit motif is insufficient to infer the underlying learning algorithm. Both BP and TL can exist on the same structures but carry out different computational roles.
Thus, BP and TL can be viewed less as separate mechanisms and more as endpoints on a continuum of learning frameworks unified by a shared objective of reducing output error through local synaptic updates guided by feedback (Aceituno et al., 2024). However, given the challenges above, theoretical modeling alone cannot definitively determine whether the brain implements BP, TL, or some hybrid.

2.6 Conclusion

This review has presented several theoretical paradigms addressing credit assignment through synaptic plasticity to enable effective learning in multi-layered hierarchical networks. While backpropagation (BP) functions as a highly effective algorithm in artificial systems, the computational mechanisms that BP requires are widely considered biologically implausible for direct implementation in the brain. In response, various approaches have been developed to relax the strict mathematical constraints of BP, which demonstrate that learning remains feasible even under biologically motivated approximations. Based on these findings, some researchers argue that biological neural networks could approximate the essential dynamics of the BP algorithm. In contrast, a distinct class of learning models, here referred to as Target Learning (TL), has emerged, which relies on fundamentally different mechanisms than those used in BP. Notably, TL models consistently operate using two distinct computational phases, whereas BP operates in one activity state during both inference and learning.
Despite these differences, BP and TL can yield equivalent weight updates under certain conditions, leading to a convergence in the weight updates. Theoretical analysis alone remains insufficient to resolve whether the cortex implements BP, TL, or a hybrid of the two. As a result, the question of which learning paradigm more accurately reflects biological learning processes remains unanswered.

3 Inference & Learning

So far we have seen that both BP and TL are optimization strategies that adapt synaptic weights in a network toward some desired output in response to some input. BP operates under the assumption that weight updates occur in response to a given input. Formally, these updates take the general form of

$$\Delta W = r_{pre}\delta^{BP}_{post} \tag{29}$$

in which r_{pre} denotes the pre-synaptic activity and δ^{BP}_{post} denotes the backpropagated error term. In comparison, weight updates in TL occur in response to a given target activity and therefore take the general form of

$$\Delta W = r^{*}_{pre}\delta^{TL}_{post} \tag{30}$$

in which r^{*}_{pre} denotes the desired activity pattern for the entire network. The synaptic weights are then adjusted to consolidate the target state.

However, as has been previously discussed, theory alone is not sufficient to determine whether BP or TL is a better candidate for explaining how biological networks learn. Therefore, we turn to contrast the two through possible experimental observations.

More concretely put, BP-based and TL-based frameworks differ fundamentally in how they conceptualize the relationship between inference and learning. In BP and approximations, the network uses the same activity that performs inference to drive synaptic updates. The feedforward pass (inference) produces an output, for which error signals to update synaptic weights are computed directly. Equation 29 shows this phenomenon, weight updates solely rely on r_{pre}. In other words, learning relies on the same activity dynamics as inference does, and the presynaptic and postsynaptic activity during normal perception or task performance is used to compute weight gradients (Whittington and Bogacz, 2017; Lillicrap et al., 2020). Implicitly, this means that BP-algorithms should show minimal divergence between activity state in the inference versus learning phases. As a result, we should observe that the network operates in a unified phase in which the activity during learning is an extension of the inference dynamics.

By contrast, TL frameworks explicitly model inference and learning as different processes. Learning requires the network to enter a distinct activity state in which target signals are present. The network communicates these distinct activity targets backward through the network. Each neuron then adjusts its synapses to consolidate the target activity. Thus, TL algorithms explicitly depend on distinct processes for learning and inference, and this should be observable through distinct activity patterns recorded in cortex. The activity state during learning reflects a feedback-modulated pattern that communicates learning signals. Therefore, TL implies there should be a bifurcation in activity depending on whether the network is in inference mode or in learning mode. The differences between inference and learning are highlighted in figure 4.

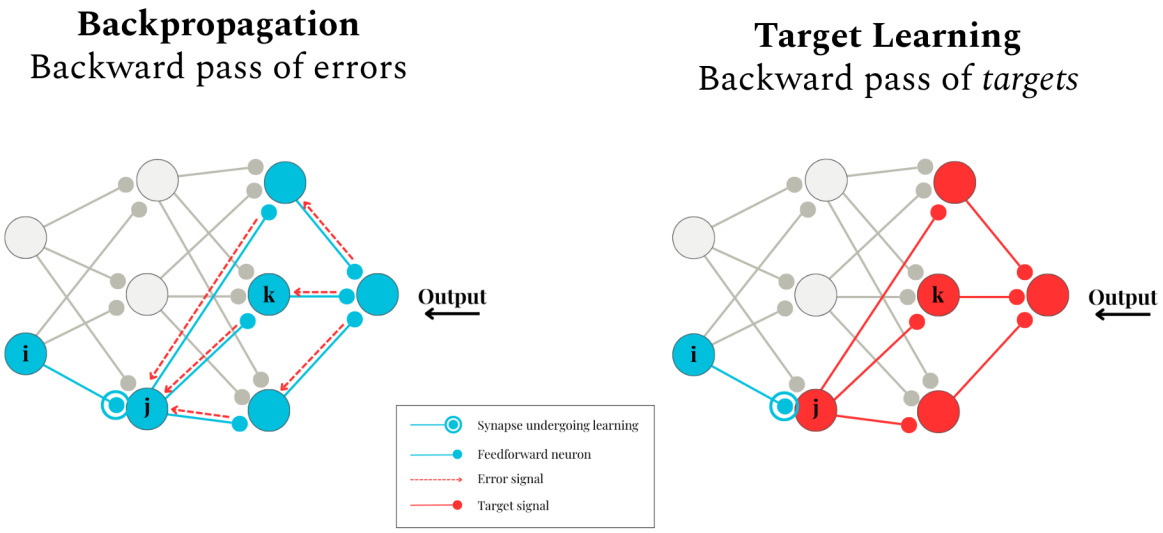


Figure 4: Schematic comparison of the difference in learning phase between BP (left) and TL (right). The system is in the same blue activity state in BP over which small errors are backpropagated, whereas for TL the system is in a different red activity state

3.1 Experimental setup

In order to test whether the brain is most likely to implement BP-like or TL-based algorithms to explain how credit assignment is addressed in hierarchical learning in cortex, it is crucial that the experimental setup clearly distinguishes between learning and inference phases. Moreover, we prefer that the data is not clouded by other important variables, so that we can clearly observe the differences between inference and learning without being convoluted by other variables that are unimportant for this purpose.

Therefore, the data set proposed by Nguyen et al. (2024) was adopted. Calcium imaging was taken from lateral visual cortex of five awake, head-fixed mice (n = 5) across daily sessions (4 or 5 sessions per mouse) while they passively viewed one of two visual stimuli (CS1 or CS2). In a trial, mice viewed CS1 or CS2 for two seconds, after which a grey screen was presented for 58 seconds to probe potential replay of the stimuli (Fig. 5A). During this 58 second ITI, we indeed observe transient spontaneous activity, moments of synchronous neural activity that occur without any stimulus being present (Fig. 5B). Recordings were taken from thousands of neurons (6,878 ± 118 neurons per session, mean ± s.e.m). However, subsequent analyses were restricted to stimulus-driven neurons[2] (1,361 ± 94 per session, mean ± s.e.m.), cells that either showed a significant preference for CS1 or CS2, or responded similarly to both.

[2]For more details on the experimental setup, preprocessing and reactivation classification pipeline we refer to Nguyen et al. (2024)

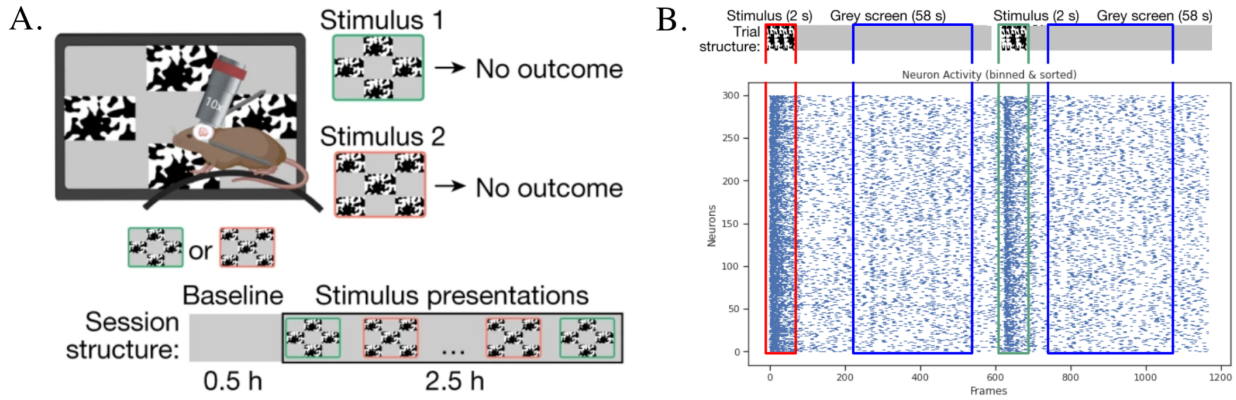


Figure 5: **A.** Graphic of the experimental setup (adapted from Nguyen et al. (2024). Lateral visual cortex is imaged while the mouse watches 1 of 2 stimuli for 2 seconds after which a grey screen is presented. The whole session lasted 3 hours, starting with 30 minutes of baseline recordings in which there was no stimulus presented. **B.** Schematic of two trials. The stimulus was presented for two seconds, after which it looked at a grey screen for 58 seconds (top panel). We observe clear stim-evoked responses in red and green in each trial, as well as transient spontaneous activity in the 58 second inter-trial-interval (ITI, bottom panel).

In the case of this dataset, we focus on four patterns of neural activity: CS1 and CS2 stimulus-evoked patterns and their reactivations, one associated to each stimulus-type. The stimulus-evoked responses are a consequence of inference in which the network simply processes the stimulus to produce output activity. By contrast, the reactivations occur without a stimulus being present and, thus, cannot be the results of inference. Such replays refers to the spontaneous reactivation of neural patterns that were active during specific experiences, often at a faster timescale than the original event. Nevertheless, it remains unclear whether the reactivations indicate that the system is operating in a learning phase.

3.2 Reactivations as Learning Signals

The phenomenon of neural replay and reactivation has been extensively studied in neuroscience across species, particularly within hippocampus and cortex. The literature highlights several functions of replay, with the most prominent focus on memory consolidation. Replay is thought to stabilize unstable memory traces and facilitates their transfer from the hippocampus to cortex for long-term storage.
Foster (2017) reviews evidence that replay during sleep supports systems memory consolidation, while awake replay contributes to memory retrieval and spatial working memory. Disruption of hippocampal sharp-wave ripples, which are associated with replay, impairs spatial learning, which higlights the essential function of replays in memory formation (Girardeau et al., 2009). Beyond consolidation, replay is implicated in planning and decision-making. Studies in rodents show that awake replay can represent future trajectories, aiding navigational planning (Foster and Wilson, 2006). In humans, Schuck and Niv (2019) found that sequential replay of nonspatial task states supports memory updating and decision-making. Replay also contributes to reinforcement learning by integrating reward-related information (Lansink et al., 2009) and skill learning by replaying experiences during rest (Gillespie et al., 2021). Moreover, ex vivo approaches of organoid neocortical microcircuit cell cultures have shown to reactivate learned patterns (Liu and Buonomano, 2025), further suggesting that replay might be an essential mechanism required for learning even in elemental lab-grown network circuits.

Whether reactivations are a static replay to support memory consolidation or serve as a learning signal is currently under heavy investigations. Most ideas lean toward the consensus that replay represents a dynamic learning signal rather than a static repetition of neural activity. Although replay involves reactivating specific patterns, replay represents an active mechanism that drives synaptic plasticity and promotes memory integration rather than functioning as mere playback to solidify learned patterns. Computational models suggest that replay emerges from networks with excitation-inhibition balance and plasticity, supporting both consolidation and retrieval (Gillespie et al., 2021). For instance, Bendor and Wilson (2012) demonstrated that replay can be biased by task-related cues and thereby imply its adaptability to behavioral demands. Deuker et al. (2013) found that spontaneous reactivations of stimulus-specific patterns during rest and sleep predict memory performance, suggesting that replay actively strengthens memory traces. Similarly, Schlichting and Preston (2014) showed that replay during rest supports learning of related content, indicating a dynamic role in memory updates. Thus, many sources have implied reactivations to be crucial for memory and learning. Here, we further seek to determine whether the reactivations are important for learning.
We hypothesize that stimulus-evoked patterns correspond to feedforward inference processes, while reactivations represent a separate, feedback-driven learning phase that interacts with inference. If both conditions, distinct reactivation patterns and their interaction with stimulus-evoked activity occur, cortical hierarchical learning operates in a TL-compatible but BP-incompatible regime. In light of this, we thus sought to determine whether:

1. the changes in the neural responses rely on a credit assignment strategy. If the data shows no signs of network-wide coordinated activity changes, then discussing whether BP or TL is more fitting would be unnecessary.

2. the reactivations drive learning. If not, we cannot distinguish between the inference and learning phases required to contrast BP with TL.

3. the system is in the same activity state during inference and learning. BP assumes that the activity states of the system should be similar in inference and learning, while TL supposes that these should be different (Fig. 4).

Another hypothesis specific to TL is that this interaction effect should be uni-directional. The feedback signal is thought to convey a learning signal that updates the feedforward weights. In other words, if the reactivations can be thought of as activity happening when the system is in the learning phase and the stim-evoked activity happening during inference, TL would expect that the reactivations influence stim-evoked activity but not vice versa.

4 Learning With Credit Assignment

A fundamental consideration in interpreting our experimental results is whether deep hierarchical learning processes are genuinely required or if simpler neural mechanisms that do not require network-wide coordination can adequately explain the observed neural activity patterns. Formally, we have previously defined a synaptic weight update to be equal to:

$$\Delta W = r_{\text{pre}} \delta_{\text{post}}, \tag{31}$$

where r_{pre} is the presynaptic term for BP or r^*_{pre} for TL and δ_{post} the postsynaptic term. Crucially, we need to check whether this δ term is operating here, or whether we can simplify by explaining the data merely with

$$\Delta W = r_{\text{pre}} g_{\text{post}} \tag{32}$$

for which g_{post} would indicate uncoordinated changes across the network, that could imply processes like Hebbian learning, synaptic pruning and homeostatic plasticity. We expect processess described by g_{post} to be present, but here we aim to establish that the data also relies on coordinated network-level changes denoted by δ_{post} simultaneously.
In practice, g_{post} would manifest as an overall decrease in global neuronal activation over repeated stimulus presentations such that most neuron activity would drop off completely. Initial widespread neural activation would become sparse, leaving a smaller, consistent subset of neurons selectively tuned to the stimulus. Such simplification of the neural code could be driven purely by local synaptic adjustments and would not require any complex coordinated learning. Under such conditions, invoking a discussion on complex deep learning strategies such as BP and TL may be premature. Thus, for the purpose of this thesis, we aim to to rule out the possibility of a simpler explanation g_{post} to sufficiently describe patterns in the data before attributing these observed patterns to more sophisticated nonlinear learning algorithms.

4.1 Methods

To test whether the neurons relied on deep hierarchical learning rather than simple uncoordinated processes, we analyzed neuronal rank ordering across early and late trials. Specifically, we computed and compared the average neuronal activity from the first five trials (early phase) against the last five trials (late phase) within each experimental session.
For explanatory purpose, we generated simulated data to visualize what we expect to see if the neurons relied solely on uncoordinated strategies (Fig. 6A & B). In these simulations, initial responses to novel stimuli would broadly activate many neurons. After repeated exposure, most neural activity would reduce, leaving a small specialized subset consistently responsive. Consequently, the neuron rankings from early to late phases would remain highly correlated, as the most active neurons initially would remain among the most responsive neurons after learning. These patterns would be inconsistent with network-wide coordinated changes, as the activity after learning could have been predicted solely from the activity before learning.

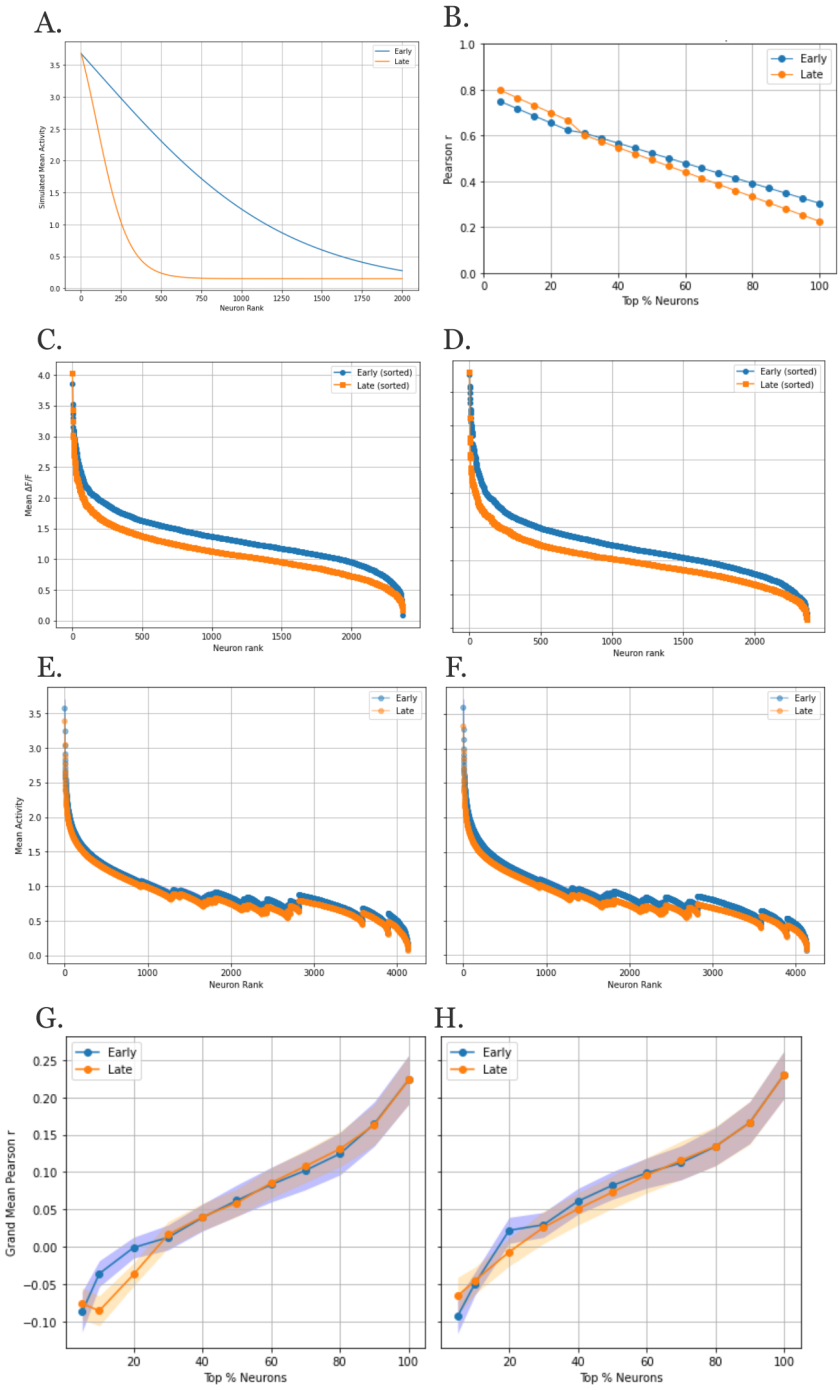


Figure 6: Credit assignment analysis. **A.** Simulated data to visualize the neuron ranks of learning without credit assignment. **B.** Simulated data to visualize expected correlation between neuron ranks of early and late trials. Without credit assignment, neuron ranks would remain correlated over learning, while most activity would drop off. **C.** Experimental results of the neuron ranks of the stim-evoked responses for CS1 averaged over early (first 5, blue) trials and late (last 5, orange) trials, one example mouse and session. **D.** Same as plot C, but for CS2. Both CS1 and CS2 show no clear drop-off in activity as a consequence of learning. **E.** CS1 neuron ranks for early (first 5, blue) trials and late (last 5, orange) trials, averaged across all mice and sessions. **F.** Same as plot E, but for CS2. **G.** Correlation between neuron ranks for early (blue) and late (orange) trials averaged across mice and session with the SEM in the shaded area for CS1. **H.** Same as plot G but for CS2.

4.2 Results

Figure 6 C - F shows the obtained results from these analyses. In figure 6 C and D we show the rank-order plot from one representative session as an example. This same pattern was consistent across mice and sessions. Plots E and F display rank-order plots averaged across multiple sessions and mice. Due to differing neuron counts across sessions, we standardized data by padding neuron counts to the session with the highest number (≈ 4000 neurons). This padding introduced noticeable waviness when averaging sorted neuronal activity rankings. However like noted earlier, all sessions showed a pattern similar to plots C and D. In plots G and H, we show the grand mean Pearson's r, averaged across sessions and mice, including the standard error of the mean (S.E.M.) in the shaded area. Pearson's r is a normalized statistic that quantifies both the strength and direction of the linear relationship between two continuous variables (Stewart, 2023). The S.E.M. was computed by the standard deviation σ across sessions divided by the square root of the number of sessions n:

$$S.E.M. = \frac{\sigma}{\sqrt{n}} \tag{33}$$

Contrary to simulated predictions, the empirical data does not show the clear activity reduction expected from simple learning mechanisms. Instead we observe no substantial drop off in activity. Moreover, when analyzing neuron rank correlation between early and late trials using Pearson's correlation (averaged across sessions and mice), we obtain negative correlations for the top $< 20\%$ neurons. This implies that neurons that were initially among the top responsive do not retain their high ranks after learning. This negative correlation indicates a reordering of neural responsiveness, that cannot be attributed to uncoordinated processes.

4.3 Conclusion

This rank-order analysis strongly suggests that relying only on simple learning mechanisms is insufficient to explain the patterns in the data. There is no substantial drop-off in neural responsiveness as a consequence of learning and the negative correlation implies a reshuffling of neuron ranks, which cannot be the result of simple learning mechanisms. Thus, we conclude that ΔW is reliant on a δ signal and that the observed changes in neural activity thus rely on a complex non-linear credit assignment strategy. We have, therefore, validated the purpose of analyzing whether BP or TL is a better candidate to explain cortical hierarchical learning.

5 Reactivations Drive Learning

Having established the need for a deep learning algorithm, we next want to test whether these reactivations are actually involved in learning, or whether they resemble something else. For this we hypothesized that the reactivations should be predictive of changes in stim-evoked activity if they are relevant to learning. In other words, if the reactivations are influencing the changes in the stim-evoked activity, they should be predictive of these changes.

5.1 Methods

To test the hypothesis that reactivations play a prescriptive role in shaping stimulus-evoked activity, we trained several logistic regression models for each session and each mouse with a variety of predictors and compared their performance.
For every consecutive stim-evoked response in each category (CS1 or CS2), we computed whether a neuron increased (labeled +1) or decreased (labeled -1) in activity w.r.t. the previous stim-evoked response. This was the target vector we tried predicting with a different set of features. For the set of features, we included the following:

1. Stim-evoked
2. Stim-evoked & Stim-evoked
3. Stim-evoked & Reactivation
4. Reactivation & Reactivation
5. Reactivation

For the first variable, we take just the vector of stim-evoked activity to predict whether a neuron will in- or decrease in the next response. For the second predictor, we substract the consecutive stim-evoked responses and use this vector to predict changes in the subsequent stim-evoked response. Including past information, therefore, contextualizes the subsequent stim-evoked response in time. Third, we introduce the reactivations and substract a given reactivation within a trial from the corresponding stim-evoked response. Fourth, we substract consecutive reactivation responses and lastly we use only the reactivation activity vector to predict changes in the subsequent stim-evoked response.
Across all sessions and mice, there were fewer classified reactivations than there were stim-evoked responses. To account for this class imbalance, we computed a target only for trials when at least one reactivation was classified in between, so that the cross-validated test set across all five folds had a standardized dimension of $(\#reactivations \times \#neurons)$.
More concretely, this analysis is directed to test the Granger causal dependence of stim-evoked plasticity on the reactivation activity. Formally, let $Y(t+1)$ denote the target, the change in a neuron's subsequent stim-evoked response (increase or decrease) w.r.t. the previous stim-evoked response and let $I(t)$ denote the full information set available at time t that includes both stim-evoked and reactivation activity. Let $I_{-R}(t)$ denote the same set excluding reactivation activity. Stim-evoked plasticity is Granger causally dependent on the reactivations if:

$$P[Y(t+1) \in A \mid I(t)] \neq P[Y(t+1) \in A \mid I_{-R}(t)] \tag{34}$$

Should the inclusion of reactivation activity shift the probability distribution of upcoming plasticity outcomes beyond what is explained by prior stimulus-evoked responses alone, we can conclude

that reactivations Granger-causally influence future stimulus-evoked activity. Granger causality thus quantifies whether past activity of one neural signal improves prediction of the future activity of another signal, but it does not establish mechanistic causation (Berzuini et al., 2012). Nonetheless, given the available data, it is the strongest type of causal dependence we can test, moving beyond correlation-based analyses.

We next examined whether feature predictiveness persisted across multiple trials. Rather than forecasting only the immediate next stimulus-evoked response, we attempted to predict responses one to five trials ahead. For lag 1, we predicted the response following the next one, for lag 2, the response two trials later, and so on. Because a genuine learning signal should inform only the very next stimulus response, we expect rapid decline in predictive accuracy even at the first lag.

Furthermore, we also test whether we can predict reactivation plasticity with the stim-evoked activity. TL hypothesizes that although the feedforward stim-evoked plasticity should be Granger causally dependent on the reactivations, this relationship is uni-directional. In other words, the reactivations should not be causally dependent on the stim-evoked activity, and the stim-evoked responses should, therefore, not have any predictive power over reactivation plasticity.

5.2 Results

We show the accuracy results averaged across sessions and mice for predicting the stimulus-evoked activity in the left plot of figure 7 and the accuracy results for predicting reactivation activity on the right. We averaged the accuracies for each category over sessions and mice and show the S.E.M. (eq. 33). At random we expect an accuracy of 0.5.

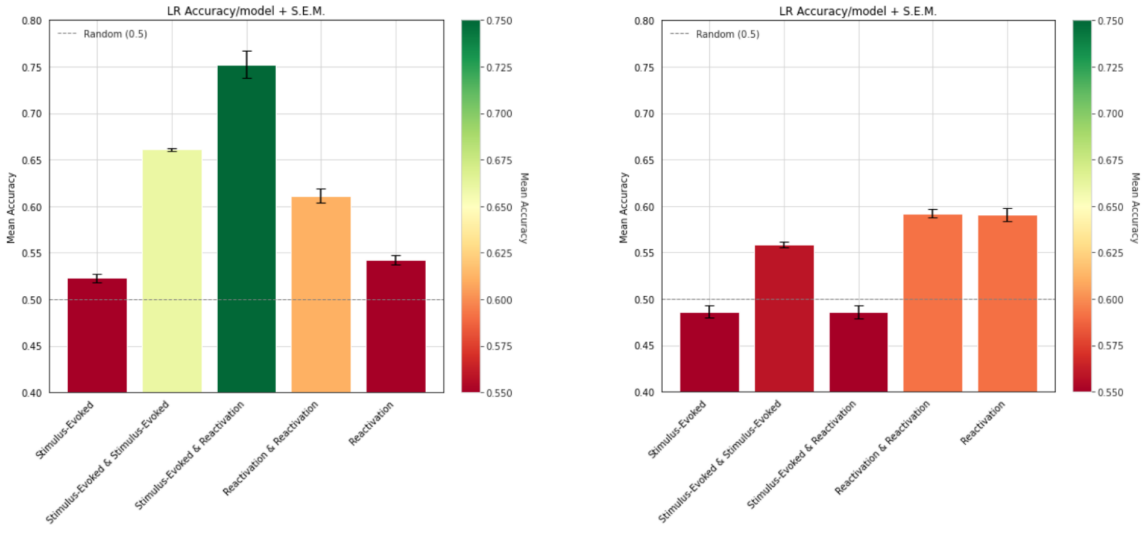


Figure 7: Logistic regression accuracy results for each feature-set, averaged across sessions and mice. **Left.** Predicting sign of change for stimulus-evoked responses. **Right.** Predicting sign of change for reactivation responses.

For the models predicting stim-evoked activity (Fig. 7, left), we observe that the model-types that rely on the single feature information of stimulus-evoked or reactivation variables do not significantly perform above chance levels. When relying only on past information of stim-evoked activity, we

obtain an accuracy of 0.66, performing significantly (section 5.3) above chance. Nonetheless, when introducing the reactivation variable to model, we obtain significantly (section 5.3) better accuracies (0.75) compared to the second best performing model that relies on past information of the stim-evoked activity. Interestingly, even when including only reactivations and thus having no access to past stim-evoked activity information, we can predict significantly above chance levels (section 5.3) with an accuracy of 0.61. We show the lag analysis for the two best model-types (Stim-evoked & Stim-evoked and Stim-evoked & Reactivation) which tests whether the learning signal is only predictive of the subsequent stim-evoked response and the predictiveness does not carry over in supplementary figure 16. Indeed, we find that the accuracy immediately drops of when predicting stim-evoked responses more time-steps in the future for the model including reactivations, whereas accuracies for the model reliant on previous stim-evoked responses remain constant. These results imply that indeed the learning signal is only relevant for the subsequent stim-evoked response and not to those that happen later in time.

For the models predicting reactivation activity (Fig. 7, right), we find that none of the models that include stim-evoked activity perform significantly better than random. As expected, the models that rely on past reactivation information do best. However, none of the models perform well, failing to achieve an accuracy above 0.6. Surprisingly, there is no difference in accuracy when including past sequential reactivation activity or use just the single previous activity vector. These findings indicate that the stim-evoked activity is not predictive of changes in reactivation responses.

5.3 Statistical Testing

To assess whether introducing reactivation information significantly improves predictive power over stim-evoked plasticity, we ran the following statistical analysis. First, we examine how model accuracy varies across individual mice and recording sessions (Supplementaries figure 14 and 15). This confirms that averaging performance over both mice and sessions does not introduce any unexpected biases or artifacts.

To compare the differences in accuracies between model performances, we first computed the balanced accuracy scores per neuron given our predicted and target per-neuron values. The balanced accuracies control for class imbalances by averaging the recall of each class. We then assessed normality of the difference in balanced accuracy between Model A (Stim-evoked & Stim-evoked) and Model B (Stim-evoked & Reactivation) per session (Shapiro-Wilk $W = 0.9411, p = 0.1899$) and across sessions per mice (Shapiro-Wilk $W = 0.9500, p = 0.7371$). Given that both per-session and per-mice balanced accuracies were non-significant for non-normality, we proceeded with paired t-tests. Across sessions, model B is significantly ($t(23) = 13.2043, p = 6.204^{-12}$ more accurate than model A. Across mice, model A also significantly outperforms model B ($t(5) = 9.2531, p = 7.5843^{-4}$). We performed two comparisons, session-level and mouse-level paired t-tests. To control the family-wise error rate, we applied a Bonferroni correction ($\alpha = 0.025$). Both session-level ($p = 6.204^{-12}$) and mouse-level ($p = 7.5843^{-4}$) results remain significant under this corrected threshold. Moreover, fitting a linear mixed-effects model with mice as random intercepts confirms that model B significantly outperforms model A (Coef = +0.281, $SE = 0.005, z = 55.130, p < 0.001$), indicating a 28% increase in balanced accuracy compared to model A.

To provide a more robust and comprehensive analysis, we also compared differences in area under the ROC curve (AUC) of prediction probabilities for both models. For all of the sessions, the auc of model B was better than model A (Fig. 8, left). We computed DeLong's tests for each neuron as it accounts for the correlated nature of AUC estimates from nested models on the same data (DeLong et al., 1988). We computed the percentage of neurons (averaged across sessions) that had a significantly (DeLong

< 0.05) better AUC in model B compared to model A. On average and across mice this percentage ranged between 20% of significant per-neuron improvement for mouse NN28 (Fig. 8, middle), to 65% percent for mouse NN11 (Fig. 8, right). Again, we test for normality on the differences between AUC per session (Shapiro-Wilk $W = 0.9586, p = 0.4350$) and per mouse (Shapiro-Wilk $W = 0.8510, p = 0.1977$). Given that both the per-session and per-mouse differences fail the non-normal distribution significance test, we continue with one-sample paired t-tests. The AUC of model B is significantly better on a per-session level ($t(23) = 15.655, p = 2.072^{-13}$) and also on a per-mouse level ($t(5) = 12.086, p = 2.69^{-4}$).

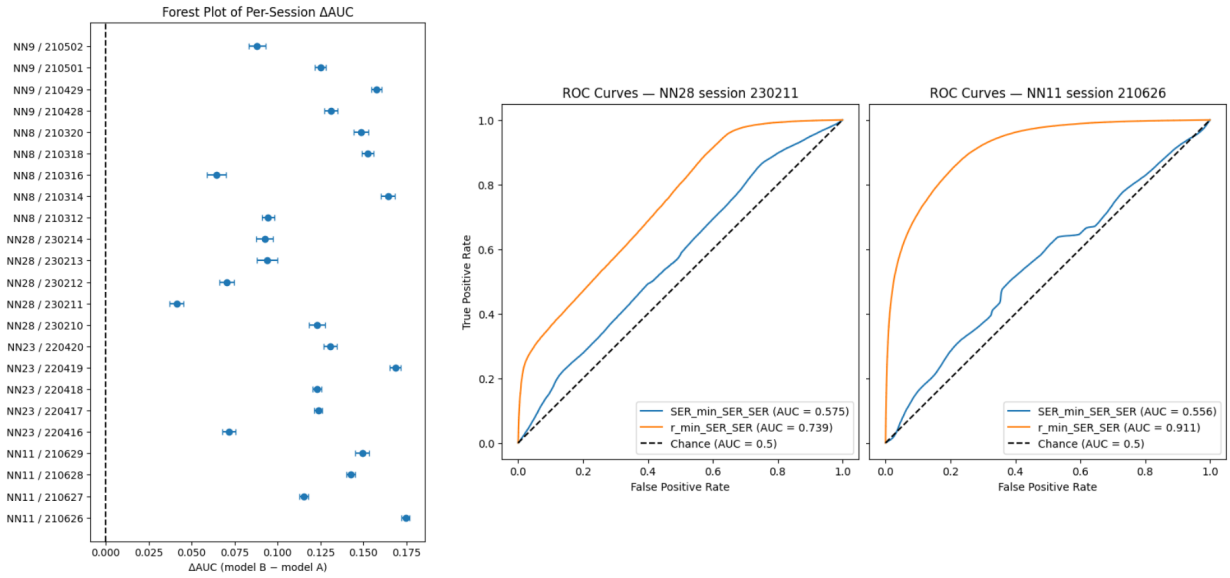


Figure 8: Statistical Results for model comparisons. **Left.** Difference in AUC of the model with reactivations versus without reactivations per mouse and per session. **Middle.** Decision curve of the model with reacticications (orange) and without reactivations (blue) for the mouse-session with the least difference in performance. **Right.** Same as the middle panel, but for the mouse-session with the largest difference in performance.

Given these statistical results we conclude that model B (including reactivations) is significantly better across mice and sessions and that, therefore, the stim-evoked plasticity is Granger causally dependent on the reactivations.

5.4 Conclusion

Taken together, we conclude that the stim-evoked plasticity is Granger causally dependent on the reactivations, but not vice versa. Formally, we successfully improve the predictive power over the stim-evoked activity $Y(t+1)$ when including the reactivation variables. Consequently, we show that the reactivations drive learning, thereby answering research question 2, and allowing use to differentiate between inference and learning phases.

6 Activity Dynamics Inference and Learning

Now that we have confirmed that the reactivations drive learning, we can examine whether cortical hierarchical learning is more compatible with BP or TL. We proposed that the main experimentally testable difference between the two learning algorithms is that they hold different assumptions on the activity dynamics across inference and learning. As we have seen before (Fig. 4), BP assumes that the system is in the same activity state during inference and learning, whereas TL supposes these states to be different.

6.1 Methods

To test these assumptions, we measure the alignment between the stim-evoked responses and the reactivations in a given trial. Concretely, we compute angle between these vectors as their cosine similarity (cossim):

$$\text{cossim}(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \, \|\mathbf{b}\|} \tag{35}$$

If the cossim between the stim-evoked responses and their reactivations is high, we are operating in the BP regime. If the cossim between the stim-evoked responses and their reactivations is low, one of the core assumptions of BP is violated. However, this would be compatible with TL, which assumes that inference and learning should display distinct activity states. The simulated hypotheses are shown in the right figure 9

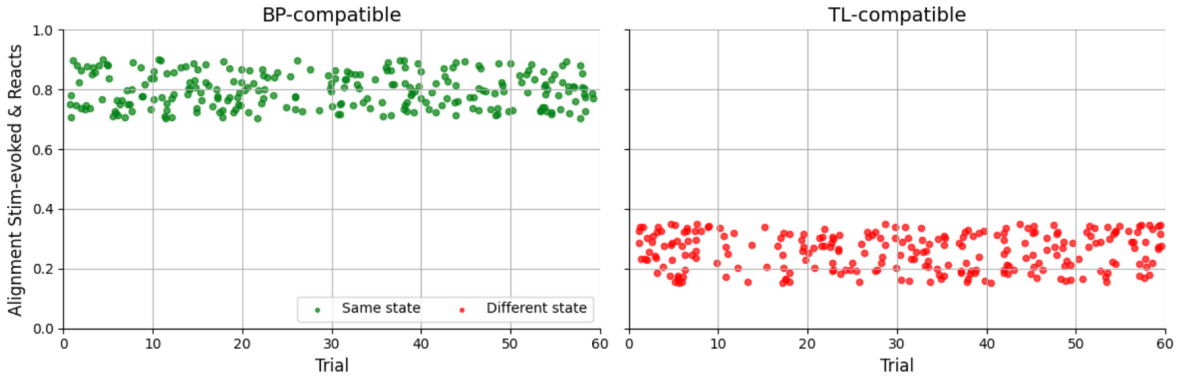


Figure 9: Simulated data to visualize BP (left), where stim-evoked and reactivations responses are aligned well (green), and TL (right), where stim-evoked and reactivations responses are not aligned well (red), hypotheses on the alignment between activity states of inference and learning.

6.2 Results

We computed the alignment between inference and learning as the cosine similarity (Eq. 35) of a stim-evoked response and the corresponding reactivation that occured in that trial. Again, for BP we expect to see high cossim between the stim-evoked responses and the reactivations. For TL we expect to see low cossim between the stim-evoked responses and the reactivations. In figure 10, we show these results averaged across sessions and mice together with the S.E.M (eq. 33).

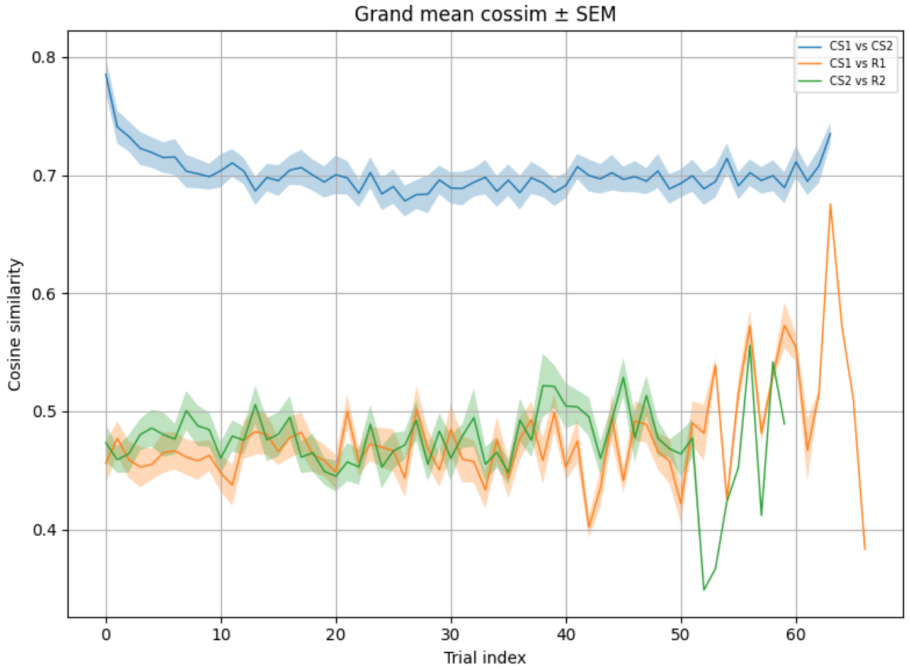


Figure 10: Cosine similarity of a stim-evoked response and a corresponding reactivation that occured within that trial, averaged across days and mice with the S.E.M. in the shaded area for each condition. Blue: cossim of the two stim-evoked response types. Orange: cossim of the stim-type 1 evoked responses (CS1) with its reactivations. Green: cossim of the stim-type 2 evoked responses (CS2) with its reactivations.

The green and orange lines show the cosine similarity between the stim-evoked responses and their corresponding reactivations. The blue line shows the cosine similarity between the two stim-evoked responses. The two stim-evoked responses are more aligned than the stim-evoked responses and their reactivations. Moreover, the alignment between the stim-evoked responses and the reactivations is similar across conditions. We show the cosine similarities between the other conditions (CS1 vs. R2, CS2 vs. R1 and R1 vs. R2) in the supplementaries 17, which indicates that the distances between the cross-conditions (CS1 vs. R2 and CS2 vs. R1) is remarkably similar to the normal condition shown here in orange and green in figure 10.

6.3 Statistical Testing

We computed the area under the curve (AUC) for the cosine similarity for each of the three conditions for each session per mouse and compared their differences to conclude whether the system's activity state is significantly different in inference and learning phase.
We fit a mixed linear regression model on the AUC, accounting for mice as random effects since we want to generalize across conditions. We find that, indeed, both CS1 vs R1 (orange) and CS2 vs R2 (green) have a significantly lower ($z(orange) = -16.051, p(orange) < .001, z(green) = -17.481, p(green) < 0.001$) AUC compared to the cossim between the stim-evoked responses (blue). Furthermore, we compute the Tukey HSD Pairwise Comparisons (Bobbitt, 2020) to test every pair of conditions, controlling for the family-wise error rate. We again find that both orange and green curve are indeed significantly different from the blue stim-evoked cossim curves

showed that the orange curve (CS1 vs R1) had a significantly lower cossim than the blue curve (CS1 vs CS2) ($Mdiff = –26.82, p < .001$), and the green curve (CS2 vs R2) likewise differed from the blue curve ($Mdiff = –29.21, p < .001$). We show the distribution of AUC across sessions for each condition in supplementaries figure 18.

6.4 Conclusion

Given these insights, we conclude that the stim-evoked responses and reactivations are not well aligned and thus, that the system is in a significantly different activity state during inference and learning phases. This finding violates one of the core assumptions of BP that supposes the system should be in the same activity state during learning and inference.

7 General Learning Signal

To get a better understanding of the stim-evoked and reactivation responses we ran several dimensionality reductions. Here, we focus the discussion on one of these techniques, the T-distributed Stochastic Neighbor Embedding (t-SNE) algorithm.

7.1 t-SNE

This algorithm is a nonlinear dimensionality-reduction technique that seeks to preserve local structure (Belkina et al., 2019). In other words, points in the data that lie close together in high dimensionality should remain close together in lower dimensions as well. In high dimensionality we compute the distance between every pair of points i, j, and convert it into a a probability $p_{j|i}$ that j is a neighbor of i. If the two points are close, probability is high, but if they are far apart, the probability is near zero. With the perplexity we tune how wide each point's neighborhood is allowed to be, usually between 5 - 50. Here we have finetuned perplexity to 30. The probabilities q_{ij} are computed similarly on how close the points end up in lower dimensionality. Then using KL-divergence, these two probabilities are matched trained to be as close as possible, thus preserving the local distances of high dimensional points in lower dimensions. We plot the results for the t-SNE trained on all data per mouse across sessions in figure 11.

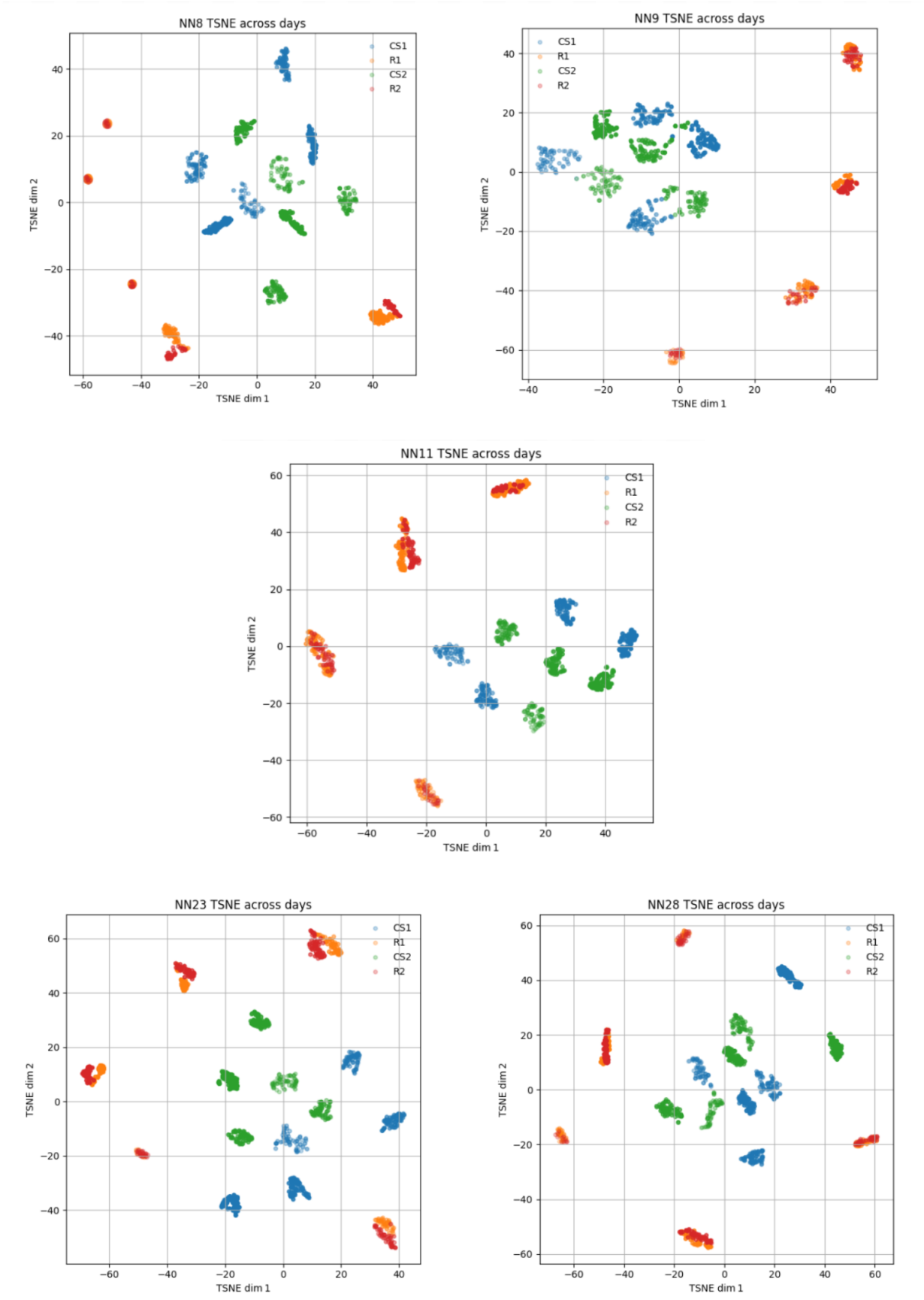


Figure 11: t-SNE dimensionality reduction plots for each mouse across its sessions. In blue and green we see the stim-evoked responses and in red and orange the reactivations. Sessions are graded by opacity, with early sessions being more transparent than later ones.

Figure 11 shows that each of the five mice, the stim-evoked patterns form distinct representations, whereas the reactivations cluster together for each session. We can count the amount of sessions by the clusters of reactivations (4 or 5) and the stim-evoked clusters will be exactly double. In other words, in lower dimensional space, there does not appear to be a clear distinction in the neural patterns that are currently classified as belonging to one of the two stimuli. Aside from supporting our conclusion of the previous section that the system is in a different activity state across inference and learning, we raised another question. Given that the reactivations occupy the same low dimensional space in every session, we wondered whether they convey a general learning signal, rather than a stimulus specific signal.

7.2 Cross-category Prediction

To test the hypothesis that the reactivations are stimulus-agnostic, we ran the same analysis as discussed in section 5, but instead swapped the stimulus-evoked responses. We thus tried predicting the sign of change of a neuron's stimulus-evoked response using the reactivations that are classified as belonging to the other category. We used the same set of features as before (described in section 5.1). We show the accuracy results for this cross-category prediction in figure 12A.

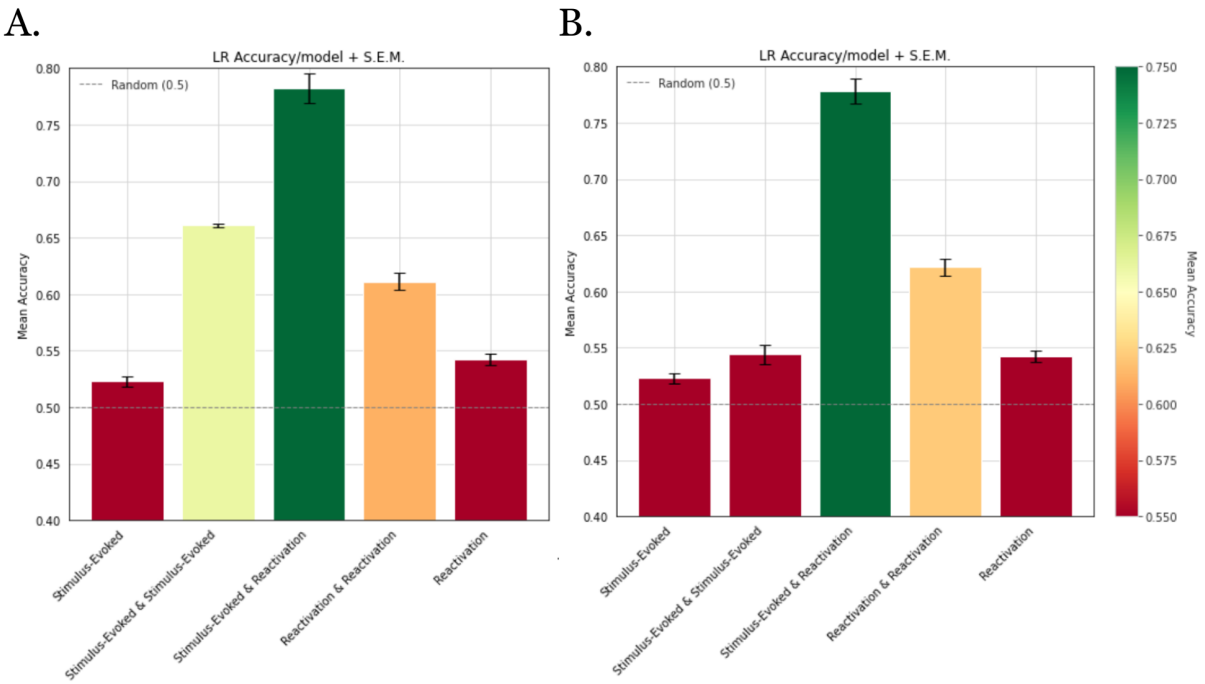


Figure 12: Logistic regression results for predicting sign of changes in the stim-evoked responses for the cross-category class (A.) and the no-category class (B.)

Interestingly enough, for the stim-evoked and reactivation model-type, we obtain prediction results of 78% accuracy, which is significantly (Sec. 7.4) better than the original implementation. In other words, reactivations currently classified as belongning to one category carry predictive information about the changes in activity for the other category.
This cross-decoding analysis supports the idea that reactivations currently classified as stimulus-type 1 carry information about changes in stimulus-evoked response type 2 and vice versa. Therefore,

we tested what would happen if we got rid of the different categories all together, and focused on predicting the changes in consecutive stim-evoked responses, regardless of whether they belonged to stim-type 1 or stim-type 2. The obtained accuracy results are shown in figure 12B. Both of these versions significantly (Sec. 7.4) outperform the normal model from section 5.

7.3 Decoding Reactivation Types

We trained a linear classifier per session per mouse to decode whether a reactivation belonged to stimulus type 1 or type 2. We show the accuracy results averaged across sessions per mouse for the linear classifiers in figure 13.

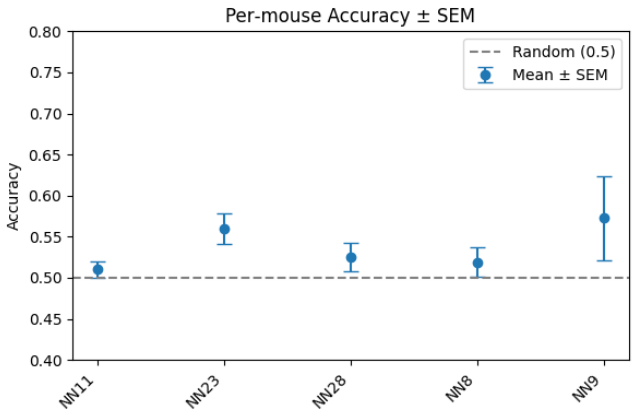


Figure 13: Accuracy scores for the linear classification of reactivations into category one or two, shown per mouse, averaged across sessions with the SEM.

We find that only for mouse 23 and mouse 9, we can decode the reactivation into the categories as they were originally classified as above chance level. However, for none of the mouse this accuracy is statistically significant above random levels.
This indicates that the reactivation representation cannot be separated linearly, thereby further supporting the idea that the classification of the reactivation as being of a specific stimulus type is unjustified in this context.

7.4 Statistical Testing

We repeat the same statistical analysis as discussed in section 5.3 to test whether there are significant differences of predictive power between each of the models (normal, cross-category, and no condition) with the Stim-evoked & Reactivation predictors. Again, we test for non-normality on the differences of balanced accuracy scores on a session-level and a mouse-level. We find that none of the models pass the non-normality Shapiro-Wilk test and we, therefore, move forward with one-sample t-tests on the balanced accuracy differences. We find that the cross-condition is significantly more accurate on both session-level ($t(23) = 4.153, p = 4.153^{-4}$) and mouse-level ($t(5) = 7.806, p = 1.453^{-3}$). Moreover, the no condition model also significantly outperforms the normal model on both the session-level ($t(23) = 4.165, p = 4.03^{-4}$) and slightly significantly on the mouse-level ($t(5) = 2.790, p = 4.931^{-2}$). Although, the no-condition model slightly significantly outperforms the cross-condition model on a mouse-level $t(5) = 3.004, p = 3.979^{-2}$), there is no significant difference on the session-level ($t(23) = 1.366, p = 0.186$). Moreover, fitting a GLM on the balanced accuracy scores confirms these findings.

Again we performed two comparisons (session-level and mouse-level paired t-tests) and to control the family-wise error rate, we applied a Bonferroni correction ($\alpha = 0.025$). Under this stricter threshold, we find that the cross-condition model still has significantly better accuracy scores than the normal model, but the no condition model only shows significance across sessions, not across mice. Moreover, under this restricted threshold, there is no significant difference between the cross-condition and no condition models. Furthermore, fitting a GLM with mice as random effects confirms these findings, where both the models significantly increases prediction accuracy (Cross: Coef = +0.048, $SE = 0.011, z = 4.552, p < 0.001$; No-cond: Coef = +0.027, $SE = 0.011, z = 2.556, p = 0.011$) compared to the normal model, but there are no significant increases in predictive power between the cross-condition no-condition models (Coef = +0.02, $SE = 0.011, z = 2.012, p = 0.054$).
We, therefore, conclude that both the models used for testing the general learning signal hypothesis are significantly better at predicting stimulus-evoked plasticity than the normal model used before was, but there is not a significant difference in accuracy between the cross- and no-condition models.

7.5 Conclusion

These results demonstrate that the predictive value of reactivations generalizes across different conditions, supporting the hypothesis that reactivations convey a general learning signal rather than stimulus-specific information. In other words, the reactivations may teach the network a broader functional rule rather than assigning explicit labels. Specifically, as observed in the t-SNE plots, the reactivations appear to guide the formation of distinct neural clusters for stimuli. However, rather than explicitly defining how each stimulus representation should individually appear, the reactivations may instruct the network that representations of distinct stimuli should remain clearly separated from each other.

8 Discussion

The findings from this thesis suggest that cortical hierarchical learning violates one of the core assumptions of the backpropagation algorithm and concludes that target learning (TL) provides better explanations of the learning dynamics. This work focuses on experimentally distinguishing between competing theories of cortical learning and approaches the question through concrete, data-driven validation.

However, a fundamental limitation of the experimental setup is the absence of behavioral measures or reward signals. Without any indication of task performance improvement, determining whether we can attribute the changes in neural responses to the process of learning remains debatable. Nonetheless, the animal does form distinctly clustered representations of the stimuli. If this phenomenon genuinely constitutes learning or results from habituation or some other neural adaptation process remains unclear. Future research should explicitly incorporate behavioral measures that indisputably defines learning. This would enable a more definitive assessment of whether target learning or backpropagation aligns more with cortical hierarchical learning.

We hypothesized that the reactivations might represent a general learning signal. To highlight this pattern of interest, we presented the t-SNE plots in section 7. However, we acknowledge that t-SNE visualizations alone are difficult to interpret due to the method's inherent complexity and sensitivity to parameter selection. To robustly confirm whether reactivations indeed represent a general function rather than merely encoding stimulus-specific labels, future studies could implement an experimental setup incorporating a third stimulus condition. The cross- and no-condition Granger Causality analyses would enable stronger insights on whether the reactivations are truly stimulus agnostic. The hypothesis of a general learning signal would be reinforced if the accuracies of predicting stim-evoked plasticity remain similar across conditions. This would imply that the reactivations are characterized by their ability to drive learning independently of specific stimulus identities in this simple experimental setup.

Moreover, our analyses revealed that the stim-evoked responses and their classified reactivations show a relatively low cosine similarity, that in lower dimensionality, the reactivation-types cluster together and that the reactivations cannot be decoded into stim-types. Given these observations and the challenges in classifying replay events due to the noisy, low temporal-resolution of calcium imaging and the fast but relatively weak nature of reactivations, we question whether the original methodology of classifying reactivations might be insufficient and potentially flawed. This concern is further intensified by the fact that the classifier also finds reactivations in the baseline period, the thirty minutes of recording preceding the very first stimulus onset (Fig. 5). So, paradoxically, we currently classify reactivations of a specific stimulus before the animal has ever seen that stimulus. The authors of the original paper argue that this might be due to the stimulus representation stitching onto existing manifolds, but in light of earlier considerations we need to interpret our findings cautiously.

Furthermore, the broader literature shows considerable disagreement on how reactivations should be classified without a universally accepted standard (Terada et al., 2022; Foster, 2017; Gupta et al., 2010; Swanson et al., 2020; Lazar et al., 2021). To complicate things further, meetings have been held to discuss the definition of the term 'reactivation' (Genzel et al., 2020). Given these limitations of uncertainty and debate, our findings underscore the need for the development of a rigorous, systematic, and standardized approach for classifying reactivation events that reliably capture their neural signatures across various conditions and experimental methodologies.

Additionally, in the computational literature, almost all bio-plausible learning rules model L5 pyramidal neurons (PNs), whereas this experimental study recorded from L2/3 neurons. This is a core limitation of the study, as studies have shown that L5 and L2/3 neurons structurally and functionally differ

from each other (Larkum et al., 2007; Tjia et al., 2017). Both neuron types do exhibit compartmental structure with basal dendrites that receive feedforward and local inputs and apical dendrites that ascend toward the cortical surface to process feedback and distant inputs. However, L5 PNs are thought of as the primary output layer of the cortex and send long-range projects to other areas. They integrate sensory information and send commands to downstream targets (Kim et al., 2015). By contrast, L2/3 PNs are primarily intracortical integrators and facilitate communication with local areas (Weiler et al., 2023). Moreover, their responses to learning and experience-dependent structural changes are less pronounced compared to L5 (Tjia et al., 2017). Furthermore, activation of L5 PNs shapes activity across other superficial layers. In auditory cortex, optogenetic activation of L5 suppresses spontaneous and evoked activity in L2/3 through the acitvations of inhibitory interneurons which suggests that L5 exerts top-down control (Onodera and Kato, 2022). As a consequence, for theories of deep credit assignment and dendritic learning, L5 PNs have remained the more compelling substrate and thus both BP-approximations and TL-based algorithms are models of these neuron-types.

Given the above, it may seem counterintuitive to test theoretical learning rules that are designed for L5 PNs on experimental data from L2/3 neurons. Nonetheless, here we argue that those bio-plausible learning rules can still be justifiably tested using L2/3 experimental recordings. Both L5 and L2/3 PNs perform compartmentalized dendritic integration, receiving feedforward inputs on basal dendrites and feedback inputs on apical tufts. This architecture is the crux of many bio-plausible learning models, such as some of the ones we discussed (predictive coding, dendritic error backpropagation, DFC), which require neurons to separately handle bottom-up and top-down signals. Since L2/3 neurons share this anatomy and excitability, they can, in theory, engage in similar learning computations. Indeed, L2/3 PNs have been shown to generate dendritic spikes and integrate apical versus basal inputs in a manner similar to L5 PNs (Larkum et al., 2007). Furthermore, cortex is highly recurrent, and L5 does not compute in isolation. This bidirectional coupling means that learning-related changes in one layer can influence, and be detected in, the other. Learning is distributed across the cortical column, so we expect to see coordinated changes across layers. L5 might implement weight updates, but L2/3 will reflect the updated computations in its firing patterns. By analyzing L2/3, we effectively probe the outcome of the deep-layer credit assignment on the entire microcircuit. Given that we are looking at more large-scale phenomena as inference and learning, and not at micro-scale specific synaptic weight updates that are not possible to observe given the current technology anyway, we argue that L5 plasticity changes resonate in L2/3. Furthermore, if a learning rule is truly bio-plausible and capable of accounting for network-wide changes, it should not be a principle that only L5 PNs can exploit. Rather, we expect it to be a general principle of cortical computation that might be most pronounced in L5 but is present in some form across layers and even across species. Importantly, even within L5 there exists a variety of different neuron-types that are thought to carry out different functions (Kim et al., 2015). Thus, restricting the explanatory scope of a learning rule to a narrow subset of neurons within a single layer undermines its generalizability. Therefore, testing hypotheses across different neuron-types such as those in L2/3 ensures that we are identifying broadly applicable mechanisms instead of layer-specific or cell-type-specific characteristics.

Continuing, the reactivations are hypothesized to carry a plasticity-related learning signal for the changes in feedforward activity. Importantly, here we are observing changes in the stim-evoked activity, but are not directly observing the feedforward weights. In practice, these changes could simply reflect fluctuations in network excitation without inducing structural changes. Critically, in our dataset, while reactivations and stimulus-evoked activity may both show changes in response patterns, we lack direct markers that would confirm synaptic strengthening. In other words, we are loosely using “plasticity signal” to refer to activity increases, but without molecular or physiological evidence, this term remains speculative. Although changes in activity patterns do reflect synaptic

weight changes, we leave open the question to what extent they resemble changes in synaptic weights. In addition, here we analyzed the calcium imaging of thousands of neurons of mouse lateral visual cortex. Calcium imaging is an indirect proxy for neural spiking activity and introduces several crucial interpretational challenges when assessing neuroplasticity in biologically plausible learning rules. First, calcium indicators record intracellular calcium transients associated with bursts of spikes, not individual action potentials. Given the framerate of 10.42 Hz per second, multiple spikes within a burst blur into a single fluorescence event, erasing precise timing and frequency information that is critical to assess spike-timing-dependent plasticity (STDP) (Markram and Sakmann, 1995; Gerstner et al., 1996; Markram et al., 2012). Moreover, the preprocessing methods enforce non-negativity of the activity and introduce thresholding assumptions, which might mask inhibitory events and could result in underrepresenting the full spectrum of network dynamics. Collectively, these limitations constrain the interpretability of calcium signals for neuro-plasticity. The observed fluorescence changes may reflect general increases in population activity rather than specific spike-pair correlations that drive synaptic modifications. As a result, inferring whether biologically plausible learning rules are truly operating is speculative. Future work should try incorporating direct measures of synaptic plasticity to confirm that the reactivations carry genuine learning signals for plasticity rather than merely envoking bursts of activity.

Lastly, we have claimed cortical learning to be more compatible with TL than with BP. However, the analyses so far have mostly focused on undermining BP as a fitting model of the patterns in the data. We have not extensively tested TL-specific hypotheses. In other words, the results do not imply that the inadequacy of BP in explaining the observed patterns necessarily validates TL as the underlying mechanism. The results from this thesis alone are, therefore, not sufficient to conclusively establish that cortical learning resembles TL. However, this thesis primarily builds upon and supplements analyses already presented in Aceituno et al. (2024). When combined with these earlier findings, the collective evidence supports the conclusion that cortical learning mechanisms are compatible with TL but inconsistent with BP.

9 Conclusion

Over the course of this thesis, we set out to analyze how well two broad families of learning algorithms, one that relies on the correcting errors, while the other relies on enforcing targets, can model experimentally observed learning phenomena. We revealed that reactivations of stimulus-evoked responses causally shape future responses, which allowed us to distinguish between inference and learning phases. Moreover, we have shown that the activity that occurs in these inference and learning phases is fundamentally different, which violates one of the core assumptions of neuro-plausible BP-approximations. In combination with findings from Aceituno et al. (2024), we argue that cortical hierarchical learning is better explained by target-based mechanisms than by backpropagation-style credit assignment. Future directions aimed at developing novel bio-plausible learning algorithms could benefit from the insights of this thesis and incorporate the discovered reactivation features for more robust learning. More broadly, this work has contributed to our understanding of cortical hierarchical computation and brings us a step closer to replicating these qualities in artificial systems.

Bibliography

Aceituno, P. V., de Haan, S., Loidl, R., and Grewe, B. F. (2024). Target learning rather than backpropagation explains learning in the mammalian neocortex. *bioRxiv*.

Akrout, M., Wilson, C., Humphreys, P. C., Lillicrap, T., and Tweed, D. (2019). Using weight mirrors to improve feedback alignment. *arXiv preprint arXiv:1904.05391*.

Bartunov, S., Santoro, A., Richards, B., Marris, L., Hinton, G. E., and Lillicrap, T. (2018). Assessing the scalability of biologically-motivated deep learning algorithms and architectures. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, pages 9368–9378.

Belkina, A. C., Ciccolella, C. O., Anno, R., Halpert, R., Spidlen, J., and Snyder-Cappione, J. E. (2019). Automated optimized parameters for t-distributed stochastic neighbor embedding improve visualization and analysis of large datasets. *Nature Communications*, 10(1):5415.

Bendor, D. and Wilson, M. A. (2012). Biasing the content of hippocampal replay during sleep. *Nature Neuroscience*, 15(10):1439–1444. Epub 2012 Sep 2.

Bengio, Y. and Frasconi, P. (1993). Credit assignment through time: Alternatives to backpropagation. In Hanson, S. J., Cowan, J. D., and Giles, C. L., editors, *Advances in Neural Information Processing Systems 6 (NeurIPS 1993)*, pages 559–566. Morgan Kaufmann.

Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166.

Berzuini, C., Dawid, A. P., and Bernardinelli, L. (2012). Granger causality. In Berzuini, C., Dawid, A. P., and Bernardinelli, L., editors, *Causality: Statistical Perspectives and Applications*, Wiley Series in Probability and Statistics, chapter 22.2.3. John Wiley & Sons Ltd, 1 edition.

Bobbitt, Z. (2020). Tukey vs. bonferroni vs. scheffe: Which test should you use? https://www.statology.org/tukey-vs-bonferroni-vs-scheffe/. Accessed: 2025-06-23.

Bottou, L., Curtis, F. E., and Nocedal, J. (2018). Optimization methods for large-scale machine learning. *arXiv preprint arXiv:1606.04838*.

Crafton, B. A., Parihar, A., Gebhardt, E., and Raychowdhury, A. (2019). Direct feedback alignment with sparse connections for local learning. *Frontiers in Neuroscience*, 13:525.

Crick, F. (1989). The recent excitement about neural networks. *Nature*.

DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*, 44(3):837–845.

Deuker, L., Olligs, J., Fell, J., Kranz, T. A., Mormann, F., Montag, C., Reuter, M., Elger, C. E., and Axmacher, N. (2013). Memory consolidation by replay of stimulus-specific neural activity. *Journal of Neuroscience*, 33(49):19373–19383.

Foster, D. J. (2017). Replay comes of age. *Annual Review of Neuroscience*, 40:581–602.

Foster, D. J. and Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440:680–683.

Friston, K. (2012). The history of the future of the bayesian brain. *NeuroImage*, 62(2):1230–1233.

Genzel, L., Dragoi, G., Frank, L., Ganguly, K., de la Prida, L., Pfeiffer, B., and Robertson, E. (2020). A consensus statement: defining terms for reactivation analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1799).

Gerstner, W., Kempter, R., van Hemmen, J. L., and Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 386:76–78.

Gillespie, A. K., Astudillo Maya, D. A., Denovellis, E. L., Liu, D. F., Kastner, D. B., Coulter, M. E., Roumis, D. K., Eden, U. T., and Frank, L. M. (2021). Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron*, 109(19):3149–3163.e6. Epub 2021 Aug 26.

Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G., and Zugaro, M. B. (2009). Selective suppression of hippocampal ripples impairs spatial memory. *Nature Neuroscience*, 12(10):1222–1223. Epub 2009 Sep 13.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, 11(1):23–63.

Gulati, T., Guo, L., Ramanathan, D., Bodepudi, A., and Ganguly, K. (2017). Neural reactivations during sleep determine network credit assignment. *Nature Neuroscience*, 20(9):1277–1284.

Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S., and Redish, A. D. (2010). Hippocampal replay is not a simple function of experience. *Neuron*, 65(5):695–705.

Hochreiter, S., Bengio, Y., Frasconi, P., and Schmidhuber, J. (2001). Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. In Kremer, S. C. and Kolen, J. F., editors, *A Field Guide to Dynamical Recurrent Neural Networks*, chapter 14. IEEE Press.

Kim, E. J., Juavinett, A. L., Kyubwa, E. M., Jacobs, M. W., and Callaway, E. M. (2015). Three types of cortical layer 5 neurons that differ in brain-wide connectivity and function. *Neuron*, 88(6):1253–1267.

Lansdell, B. J., Prakash, P. R., and Kording, K. P. (2019). Learning to solve the credit assignment problem. *arXiv preprint arXiv:1906.00889*.

Lansink, C. S., Goltstein, P. M., Lankelma, J. V., McNaughton, B. L., and Pennartz, C. M. A. (2009). Hippocampus leads ventral striatum in replay of place-reward information. *PLoS Biology*, 7(8):e1000173. Epub 2009 Aug 18.

Larkum, M. (2013). A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends in Neurosciences*, 36(3):141–151.

Larkum, M. E., Waters, J., Sakmann, B., and Helmchen, F. (2007). Dendritic spikes in apical dendrites of neocortical layer 2/3 pyramidal neurons. *Journal of Neuroscience*, 27(34):8999–9008.

Launay, J., Poli, I., and Krzakala, F. (2019a). Principled training of neural networks with direct feedback alignment. *arXiv preprint arXiv:1906.04554*.

Launay, J., Poli, I., and Krzakala, F. (2019b). Principled training of neural networks with direct feedback alignment. *arXiv preprint arXiv:1906.04554*.

Lazar, A., Lewis, C., Fries, P., Singer, W., and Nikolic, D. (2021). Visual exposure enhances stimulus encoding and persistence in primary cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 118(43):e2105276118.

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*, 7:13276.

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.

Liu, B. and Buonomano, D. V. (2025). Ex vivo cortical circuits learn to predict and spontaneously replay temporal patterns. *Nature Communications*, 16:3179.

Markram, H., Gerstner, W., and Sjöström, P. J. (2012). Spike-timing-dependent plasticity: a comprehensive overview. *Frontiers in Synaptic Neuroscience*, 4:2.

Markram, H. and Sakmann, B. (1995). Action potentials propagating back into dendrites triggers changes in efficacy. In *Society for Neuroscience Abstracts*, volume 21.

Meulemans, A., Farinha, M. T., Cervera, M. R., Sacramento, J., and Grewe, B. F. (2022). Minimizing control for credit assignment with strong feedback. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 15596–15630. PMLR.

Meulemans, A., Farinha, M. T., Ordóñez, J. G., Aceituno, P. V., Sacramento, J., and Grewe, B. F. (2021). Credit assignment in neural networks through deep feedback control. In *Advances in Neural Information Processing Systems*, volume 34.

Moskovitz, T. H., Litwin-Kumar, A., and Abbott, L. F. (2018). Feedback alignment in deep convolutional networks. *arXiv preprint arXiv:1812.06488*.

Nguyen, N. D., Lutas, A., Amsalem, O., Fernando, J., Ahn, A. Y., Hakim, R., Vergara, J., McMahon, J., Dimidschstein, J., Sabatini, B. L., and Andermann, M. L. (2024). Cortical reactivations predict future sensory responses. *Nature*, 625(7993):110–118.

Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. Springer, 2nd edition.

Nøkland, A. (2016). Direct feedback alignment provides learning in deep neural networks. In *Neural Information Processing Systems*.

Onodera, K. and Kato, H. K. (2022). Translaminar recurrence from layer 5 suppresses superficial cortical layers. *Nature Communications*, 13:2585.

Ororbia, A., Mali, A., Kohan, A., Millidge, B., and Salvatori, T. (2024). A review of neuroscience-inspired machine learning. *arXiv preprint arXiv:2403.18929*.

Ororbia, A. G. (2023). Brain-inspired machine intelligence: A survey of neurobiologically-plausible credit assignment. *arXiv preprint arXiv:2312.09257*.

Rahnev, D. (2019). The bayesian brain: What is it and do humans have it? *Behavioral and Brain Sciences*, 42:e238.

Rao, R. P. N. and Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87.

Richards, B. A. and Lillicrap, T. P. (2019). Dendritic solutions to the credit assignment problem. *Current Opinion in Neurobiology*, 54:28–36.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323:533–536.

Rumelhart, D. E. and McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 1: Foundations*. MIT Press eBooks. MIT Press.

Sacramento, J., Costa, R. P., Bengio, Y., and Senn, W. (2018). Dendritic cortical microcircuits approximate the backpropagation algorithm. In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, pages 8735–8746. Curran Associates, Inc.

Scellier, B. and Bengio, Y. (2016). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *arXiv preprint arXiv:1602.05179*.

Schlichting, M. L. and Preston, A. R. (2014). Memory reactivation during rest supports upcoming learning of related content. *Proceedings of the National Academy of Sciences of the United States of America*, 111(44):15845–15850.

Schuck, N. W. and Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, 364(6447):eaaw5181.

Semedo, J. D., Jasper, A. I., Zandvakili, A., Machens, C. K., and Kohn, A. (2022). Feedforward and feedback interactions between visual cortical areas use different population activity patterns. *Nature Communications*, 13(1):1099.

Song, Y., Millidge, B., Salvatori, T., Lukasiewicz, T., Xu, Z., and Bogacz, R. (2024). Inferring neural activity before plasticity as a foundation for learning beyond backpropagation. *Nature Neuroscience*, 27:348–358.

Spratling, M. W. (2019). Fitting predictive coding to the neurophysiological data. *Brain Research*, 1720:146313.

Stewart, K. (2023). Pearson's correlation coefficient — definition, formula, & facts — britannica. https://www.britannica.com/topic/Pearsons-correlation-coefficient. Accessed: 2025-06-23.

Stolyarova, A. (2018). Solving the credit assignment problem with the prefrontal cortex. *Frontiers in Neuroscience*, 12:182.

Stork, D. G. (1989). Is backpropagation biologically plausible? In *Proceedings of the International Joint Conference on Neural Networks*, volume 2, pages 241–246, Washington, DC, USA. IEEE.

Swanson, R. A. et al. (2020). Variable specificity of memory trace reactivation during hippocampal sharp wave ripples. *Current Opinion in Behavioral Sciences*, 32:126–135.

Terada, S. et al. (2022). Adaptive stimulus selection for consolidation in the hippocampus. *Nature*, 601:240–244.

Tjia, M., Yu, X., Jammu, L. S., Lu, J., and Zuo, Y. (2017). Pyramidal neurons in different cortical layers exhibit distinct dynamics and plasticity of apical dendritic spines. *Frontiers in Neural Circuits*, 11:43.

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., and Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 108(51):20754–20759.

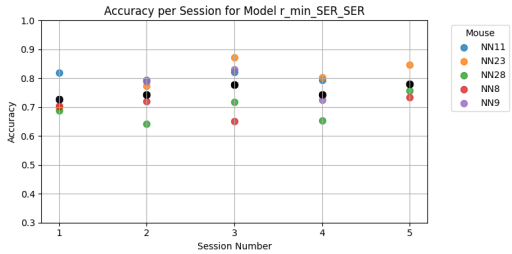
Weiler, S., Guggiana Nilo, D., Bonhoeffer, T., Hubener, M., Rose, T., and Scheuss, V. (2023). Functional and structural features of l2/3 pyramidal cells continuously covary with pial depth in mouse visual cortex. *Cerebral Cortex*, 33(7):3715–3733.

Whittington, J. C. R. 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.

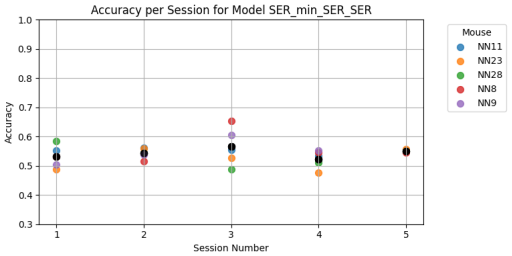
Zagha, E. (2020). Shaping the cortical landscape: Functions and mechanisms of top-down cortical feedback pathways. *Frontiers in Systems Neuroscience*.

Supplementaries

A Granger Causality

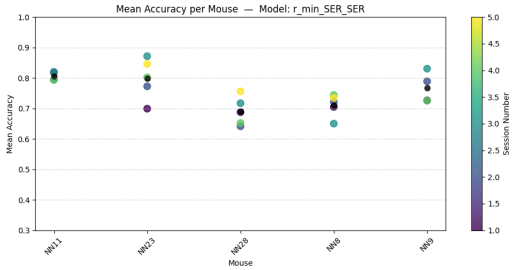


(a) Accuracy distribution per session across mice for model stim-evoked and reactivations

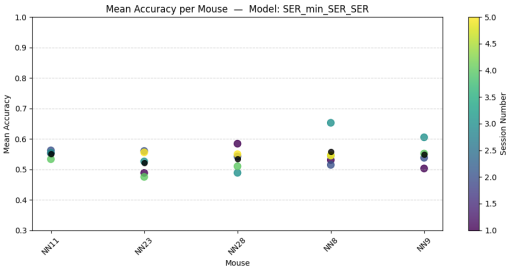


(b) Accuracy distribution per session across mice for model stim-evoked and stim-evoked

Figure 14: Session-wise accuracy distributions across all mice.



(a) Accuracy distribution across sessions per mouse for model stim-evoked and reactivations



(b) Accuracy distribution across sessions per mouse for model stim-evoked and stim-evoked

Figure 15: Mouse-wise accuracy distributions across all sessions.

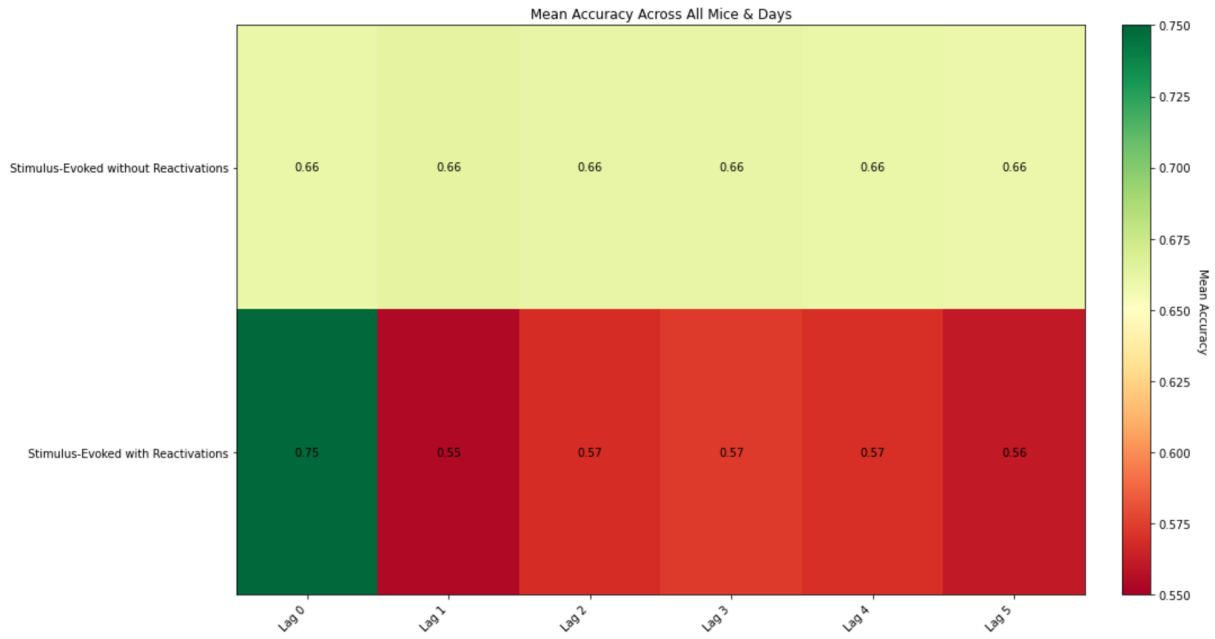


Figure 16: Lag analysis to test prediction accuracy of stim-evoked responses with 1 to 5 steps in the future. **Top**: Accuracies for model stim-evoked & stim-evoked. **Bottom**: Accuracies for model stim-evoked & reactivations.

B Cosine Similarities

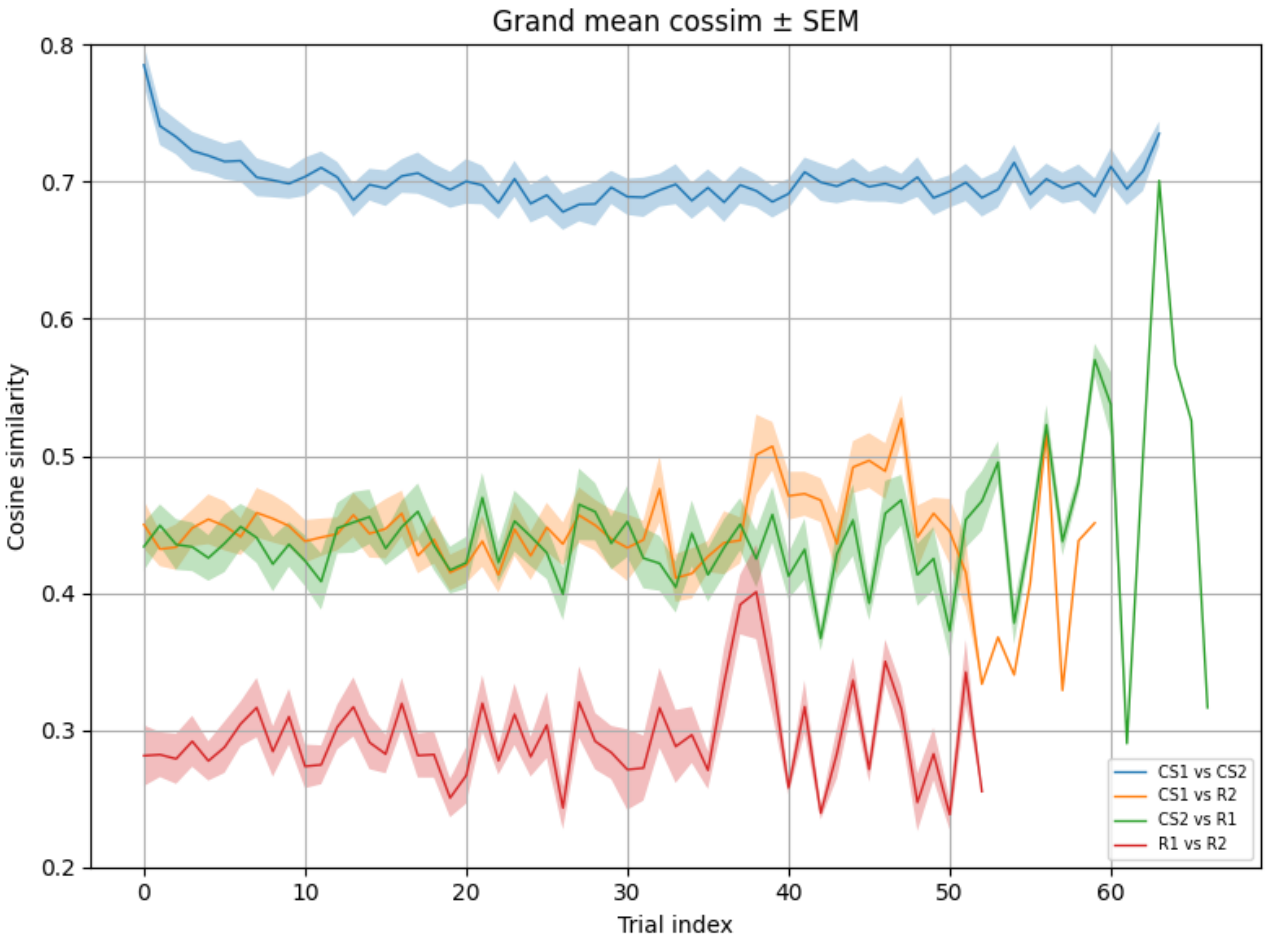


Figure 17: Cosine similarity between a stim-evoked response and a corresponding reactivation that occured within that trial, averaged across days and mice with the SEM in the shaded area for each condition. Blue: cossim of the two stim-evoked response types. Orange: cossim of the stim-type 1 evoked responses with CS2 reactivations. Green: cossim of the stim-type 2 evoked responses with CS1 reactivations. Red: cossim between the two types of reactivation responses

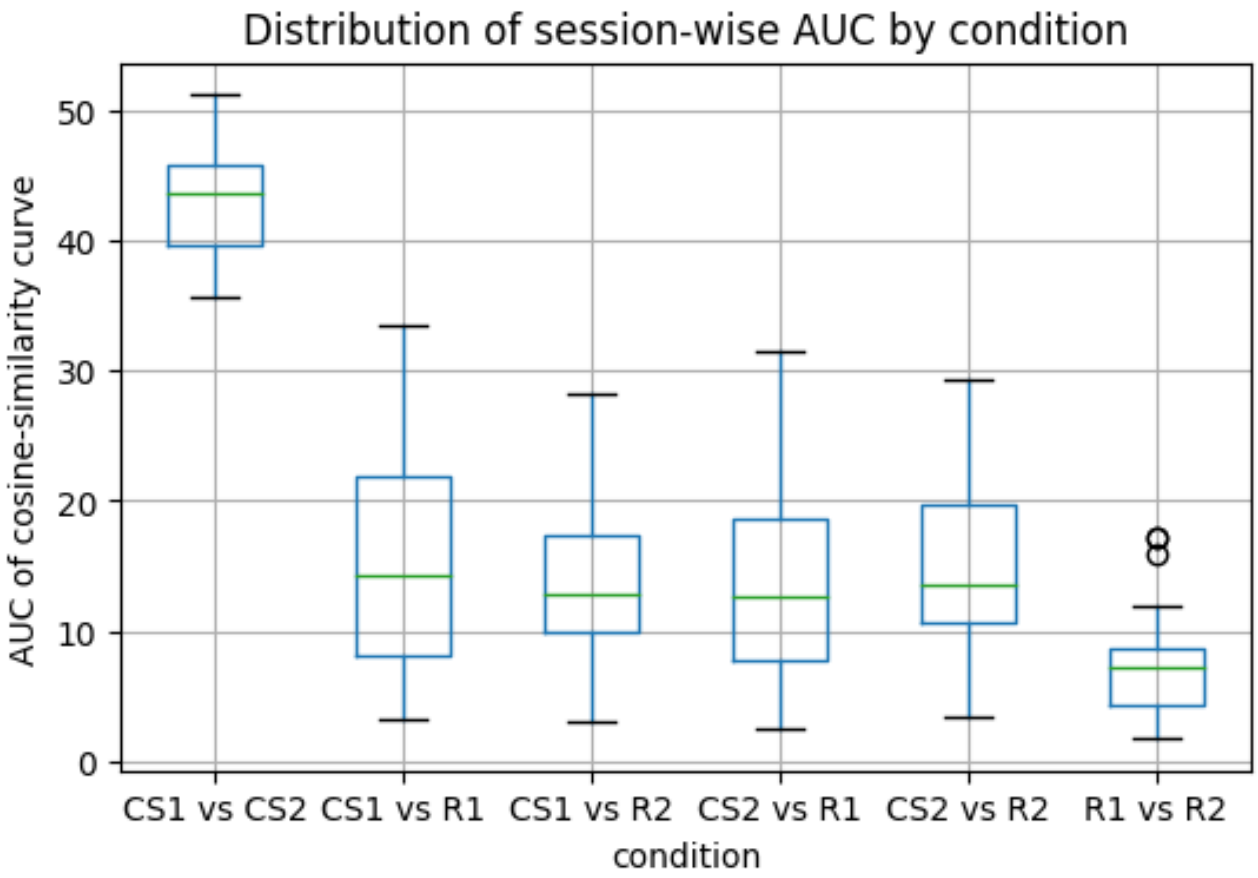


Figure 18: Distribution of AUC of all conditions.