Challenging Backpropagation: Evidence for Target Learning in the Cortex

Pau Vilimelis Aceituno[*,1,3], Sander de Haan[*,1,2,3], Reinhard Loidl[1], and Benjamin F. Grewe[1,3,†]

[1]Institute of Neuroinformatics, University of Zurich and ETH Zurich
[2]Institute of Social Ethics, University of Lucerne
[3]ETH AI Center, ETH Zurich
[†]Correspondence: bgrewe@ethz.ch
[*]Equal contribution

November 29, 2024

1 Abstract

Studies at the intersection of neuroscience and machine learning have offered new insights to explain hierarchical learning in the neocortex. Two competing hypotheses have emerged: deep learning-inspired approximations of the backpropagation algorithm, where neurons adjust synapses to minimize the error, and target learning algorithms, where neurons learn by reducing the feedback needed to achieve a desired activity. Despite decades of research and theoretical arguments supporting either possibility, there is currently no conclusive evidence for either hypothesis. We address this long-standing question by focusing on the relationship between synaptic plasticity and the somatic activity of pyramidal neurons. We first build a pyramidal neuron model integrating subcellular processes including calcium dynamics, backpropagating action potentials, and plateau potentials. Our model predicts that apical synaptic inputs drive basal synaptic plasticity through somatic depolarization caused by plateau potentials. We then test this prediction through in vitro electrophysiology experiments in which we co-stimulate apical and basal synapses to induce basal plasticity. These results allow us to derive distinct predictions for both the target learning and backpropagation hypotheses which we test on in vivo neuronal activity data from the mouse visual cortex. Our findings reveal that cortical learning is consistent with target learning, but not backpropagation, highlighting a critical discrepancy between deep learning and hierarchical learning in the neocortex.

2 Introduction

The mammalian neocortex maps complex sensory inputs into abstract hierarchical representations through the coordinated activity of large neuronal networks. However, the algorithmic principles behind learning such hierarchical representations had been debated for decades(Grossberg, 1987; Crick, 1989; Richards et al., 2019; Saxe et al., 2021), and yet remain largely unknown.

On one hand, proponents of BP argue that cortical circuits have the necessary circuits and architectures that could perform the gradient computations necessary for BP (Sacramento et al.,

2018; Scellier and Bengio, 2017; Whittington and Bogacz, 2017; Lillicrap et al., 2016, 2020). On the other, various alternative explanations of hierarchical learning have been proposed that rely on predictive principles (Rao and Ballard, 1999; Song et al., 2024), energy models (Rao and Ballard, 1999; Scellier and Bengio, 2017), or control theory (Meulemans et al., 2021, 2022b). In this work, we consolidate many of these explanations into a single alternative to BP, which we refer to as 'target learning' (TL)—a family of algorithms where learning is driven by enforcing target network activity, with synapses adapting to minimize the signal strength required to reach this target (Rao and Ballard, 1999; Meulemans et al., 2021; Song et al., 2024). Although various computational studies did relate cortical learning to either BP (Körding and König, 2001; Richards et al., 2019) or TL (Lässig et al., 2023; Song et al., 2024), they rely on theoretical arguments and do not offer conclusive biological evidence, leaving open the question of which family of algorithms best explains cortical learning.

To bridge the gap between theory and biology, researchers have proposed bio-plausible cortical circuits that implement BP and TL. However, these models differ in their assumptions and the specific experimental observations they explain (Bastos et al., 2012; Lillicrap et al., 2016; Sacramento et al., 2018; Payeur et al., 2021; Aceituno et al., 2023), thereby obscuring which family of algorithms best accounts for cortical learning. Experimentally validating these circuits is further complicated by the immense complexity of cortical networks and the fact that similar circuits can implement different algorithms with only slight modifications (Whittington and Bogacz, 2017; Scellier and Bengio, 2017; Meulemans et al., 2021, 2022a; Aceituno et al., 2023; Song et al., 2024).

Regardless of the specific circuits or algorithms that have been proposed, all models of cortical learning must make specific assumptions about how pyramidal neurons (PNs), the primary building blocks of the cortex, tune their basal dendritic synapses to relay feedforward information. Those assumptions, however, differ across models and result in different predictions (Whittington and Bogacz, 2017; Sacramento et al., 2018; Payeur et al., 2021; Meulemans et al., 2021, 2022a; Aceituno et al., 2023; Song et al., 2024). Therefore, understanding how learning occurs in PNs is essential for resolving the debate on cortical learning.

We thus focus our investigation on the dynamics and plasticity of PNs and their basal dendritic synapses. We start by constructing a PN model from the ground up, based on well-established experimental observations of cellular processes and synaptic plasticity. We then use this model to make predictions about how PNs adjust their basal synapses in response to feedforward sensory inputs and feedback learning signals, which we test using in vitro electrophysiology. Finally, we leverage our single neuron model to formulate tests on the population level, which we apply to in vivo data, evaluating BP and TL as competing hypotheses for cortical learning.

3 Results

3.1 Modeling basal synaptic plasticity and pyramidal neuron dynamics

We start by modeling plasticity in basal synapses based on the current understanding of molecular-level observations showing that synaptic calcium is one key driver of plasticity in cortical neurons (Lisman, 1989; Yang et al., 1999; Evans and Blackwell, 2015). Depending on the synaptic calcium concentration, calcium-dependent plasticity pathways based on protein kinases or phosphatases are activated, leading to long-term potentiation (LTP) or long-term depression (LTD), respectively. The concentration and temporal dynamics of synaptic calcium thereby dictate the direction and magnitude of synaptic plasticity. Individual synapses maintain their strength when calcium concentrations are below an LTD threshold, but exceeding this threshold induces LTD until a second threshold is reached, beyond which LTP is triggered, leading to a plasticity function P — outlined in Graupner

and Brunel (2012) (see **Fig. 1B**) — which we adopt in our model as,

$$\Delta w = \int P(c(t))dt \tag{1}$$

where Δw is the change in synaptic strength and $c(t)$ is the calcium concentration.

In pyramidal neurons, synaptic calcium influx is regulated via the opening of NMDA receptors (NMDARs) that are essential for plasticity (Artola and Singer, 1987; Bear et al., 1992; Markram et al., 1997; Sjöström et al., 2003). NMDARs require glutamate release from presynaptic activity as well as postsynaptic membrane depolarization (Emptage et al., 1999) to open. We thus model the dynamics of synaptic calcium through NMDARs as,

$$\tau_c \dot{c} = -c + (1 - \frac{c}{c_{\max}}) I(v_{\text{synapse}}, g_{pre}) \tag{2}$$

where $c_{\max}$ is the saturated calcium concentration dictated by the electrochemical gradient, τ_c is the calcium time constant, I corresponds to the influx of calcium through NMDARs characterized by McRory et al. (2001); Slutsky et al. (2004), g_{pre} is the presence or absence of glutamate and v_{synapse} represents the postsynaptic membrane depolarization.

The postsynaptic membrane is depolarized by multiple sources within the PN. Besides EPSPs, backpropagating action potentials (bAPs) originating from the axon hillock depolarize the postsynaptic membrane shortly after somatic APs (Stuart and Sakmann, 1994). bAPs have been put forth as driving synaptic changes in spike time-dependent plasticity (STDP) (Markram, 1997; Sjöström et al., 2010), however, more recent experimental data suggests that bAPs by themselves are insufficient to explain plasticity in physiologically relevant extracellular calcium concentrations (Inglebert et al., 2020; Chindemi et al., 2022). We therefore also consider plateau potentials, which are neuron-scale depolarization events originating from calcium influx through voltage-gated calcium channels (VGCCs) at the main bifurcation of the apical dendrite, which can depolarize the postsynaptic membrane for sustained periods (Larkum et al., 1999). Accordingly, we define our synaptic membrane model as,

$$\tau_v \dot{v}_{\text{synapse}} = -v_{\text{synapse}} + \sum_{t \in \text{EPSPs}} \kappa_{\text{EPSP}}^{\text{synapse}} * \delta(t) + \sum_{t \in \text{bAPs}} \kappa_{\text{bAP}}^{\text{synapse}} * \delta(t) + \sum_{t \in \text{Plateaus}} \kappa_{\text{plateau}}^{\text{synapse}} * \delta(t) \tag{3}$$

where τ_v corresponds to the membrane voltage time constant, $\delta(t)$ stands for the Dirac delta function at time t, the $*$ denotes the convolution operator, and κ indicates the response induced by the subscripted event in the superscripted compartment (for all model hyperparameters see **Methods 1**).

We next expand our synaptic model by adding somatic and apical-dendritic compartments. In L5 apical dendrites, the tuft region forms a distinct integrative compartment (Coogan and Burkhalter, 2004; Larkum et al., 1999, 2001; D'Souza and Burkhalter, 2017), integrating top-down inputs from higher cortical and subcortical areas (Wimmer et al., 2010). Typically, top-down inputs to the apical tuft either excite or suppress local dendritic activity but do not induce a plateau potential or trigger somatic APs on their own and single bAPs are usually insufficient to initiate a plateau potential (Grewe et al., 2010). However, when inputs to the apical dendrite coincide with a bAP induced by basal inputs, the resulting depolarization at the apical dendrite can cross a threshold to open VGCCs, after which plateau potentials initiate (Larkum et al., 1999). Once generated, these plateau potentials propagate towards the somato-axonic region, where the membrane depolarization facilitates short, high-frequency bursts of APs, also known as bAP-activated Ca^{2+} spike firing (BAC firing).

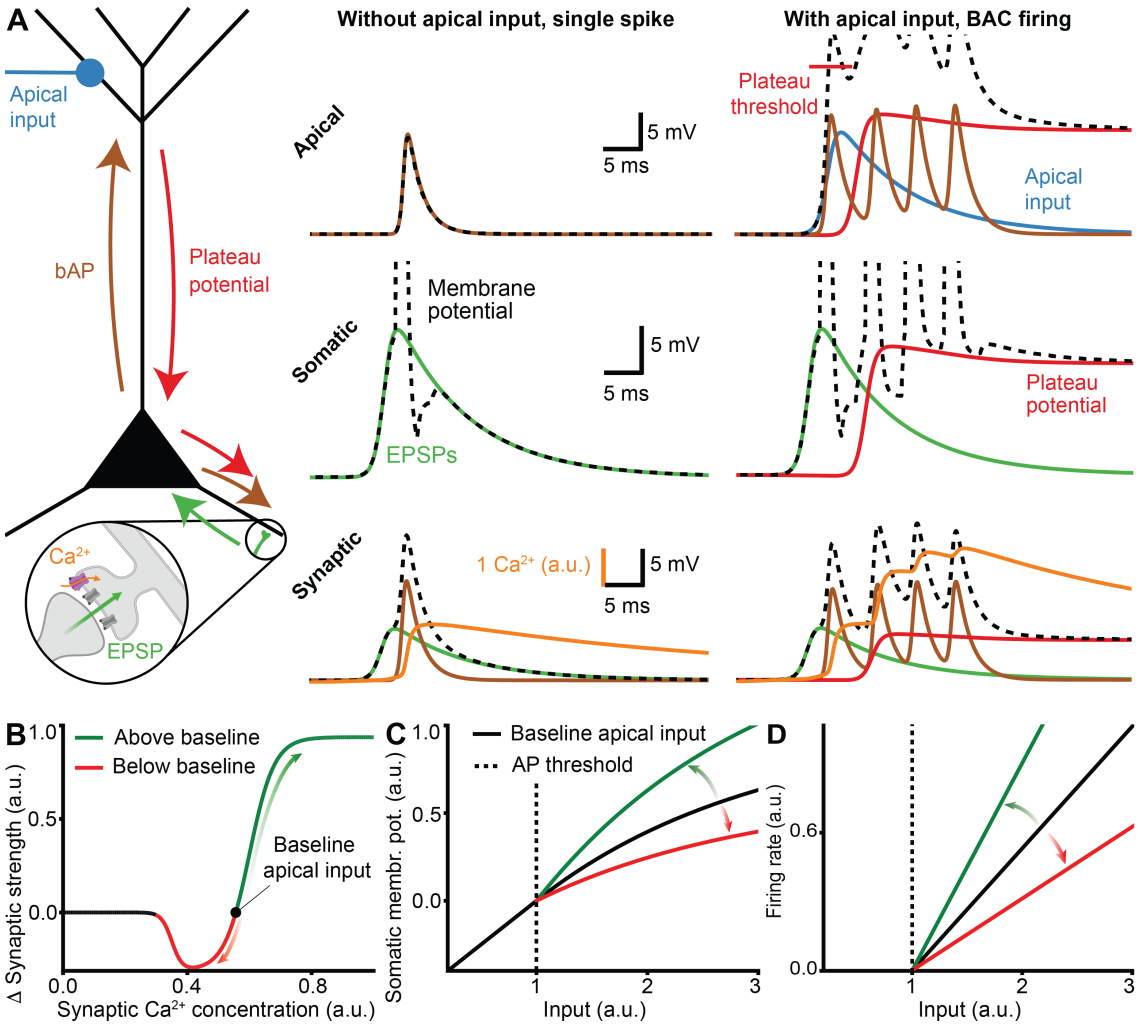

Figure 1: Modeling the integration of basal and apical inputs in PNs. (**A**) Neuron schematic (left) details all events and interactions in the model. Responses of each compartment are shown for apical, somatic, and synaptic compartments for cases without (middle) and with (right) top-down input to the apical dendrite. (**B**) Synaptic plasticity as a function of calcium concentrations. (**C**) Synaptic membrane potential in expectation as a function of basal and apical input. (**D**) Firing rate in expectation as a function of basal and apical input. For (**B-C-D**) the apical input with respect to baseline (black) has a modulating effect, with inputs above baseline leading to LTP and increased synaptic membrane potential as well as firing rate (green), and below baseline to LTD and decreased synaptic membrane potential as well as firing rate.

Accordingly, we treat somato-axonic and apical-dendritic regions as distinct integrative compartments that exchange information through bAPs and plateau potentials,

$$\tau_v \dot{v}_{\text{apical}} = -v_{\text{apical}} + \sum_{t \in \text{EPSPs}} \kappa^{\text{apical}}_{\text{EPSP}} * \delta(t) + \sum_{t \in \text{Plateaus}} \kappa^{\text{apical}}_{\text{plateau}} * \delta(t) + \sum_{t \in \text{bAPs}} \kappa^{\text{apical}}_{\text{bAP}} * \delta(t) \tag{4}$$

$$\tau_v \dot{v}_{\text{soma}} = -v_{\text{soma}} + \sum_{t \in \text{Pre}} \kappa^{\text{soma}}_{\text{pre}} * \delta(t) + \sum_{t \in \text{Plateaus}} \kappa^{\text{soma}}_{\text{plateau}} * \delta(t) + \sum_{t \in \text{APs}} \kappa^{\text{soma}}_{\text{AP}} * \delta(t) \tag{5}$$

This allows bAPs to spread from the somatic compartment to the apical and synaptic compartments and likewise the plateau potentials to spread from the apical compartment to the somatic and synaptic compartments (see **Fig. 1A**). In our model, this bidirectional link has two functions. First, the sustained depolarization from plateau potentials facilitates BAC firing. Second, bAPs and plateau potentials depolarize the basal synapses which in turn modulate plasticity through NMDAR-regulated synaptic calcium influx.

In summary, our neuron model couples apical input with both the firing rate of the neuron and synaptic plasticity through somatic and synaptic membrane depolarization (see **Fig. 1C&D** respectively). While basal EPSPs coupled with single bAPs cause brief synaptic depolarization, the quick succession of bAPs together with the plateau potential during BAC firing results in prolonged synaptic depolarization. Consequently, the basal synaptic membrane potential increases proportionally to the apical input, which in turn influences synaptic calcium influx through NMDARs when glutamate from a preceding EPSP is present (see **Fig. 1B**). As a result, our model displays single APs and moderate synaptic calcium influx in the absence of apical input, and larger synaptic calcium concentrations when basal and apical inputs coincide (see **Fig. 1A** and **Fig. S3** for an extended visual summary). This increased firing rate is consistent with experimental findings that apical input significantly modulates the firing rate of a neuron in a multiplicative manner (Larkum et al., 2004).

3.2 Apical input directs synaptic plasticity in basal synapses and affects PN activity

To test the predictions of our model that apical input modulates the firing rate of the neuron and directs the basal synaptic plasticity, we next performed in vitro whole-cell patch clamp experiments on a total of 12 L5 PNs (6 animals) from the mouse prefrontal cortex with extracellular electric stimulations of basal and apical afferents (see **Fig. 2A**).

We recorded basally-induced baseline EPSPs for 5 minutes to establish initial synaptic responses. Next, we increased the basal stimulation intensity to induce a single action potential and further adjusted apical stimulation intensity to provoke supra-threshold events that consisted of one, two, or three action potentials and prolonged somatic depolarization indicative of a plateau potential. After repeating these paired stimulation-induced supra-threshold events (PSSTs) 8 times, we again recorded basally-induced EPSPs for at least 30 minutes to measure long-term changes in EPSP amplitudes (see **Fig. 2B** and **Methods 2**). We conducted all recordings in physiological extracellular calcium concentrations (Inglebert et al., 2020).

Comparing basally-induced EPSPs before and after PSSTs, we found that some stimulation variants led to significantly increased EPSP amplitudes indicative of LTP (see **Fig. 2C-E top**, $p < 0.0001$), while others resulted in no significant difference of EPSP amplitudes (see **Fig. 2C-E bottom**, $p = 0.229$).

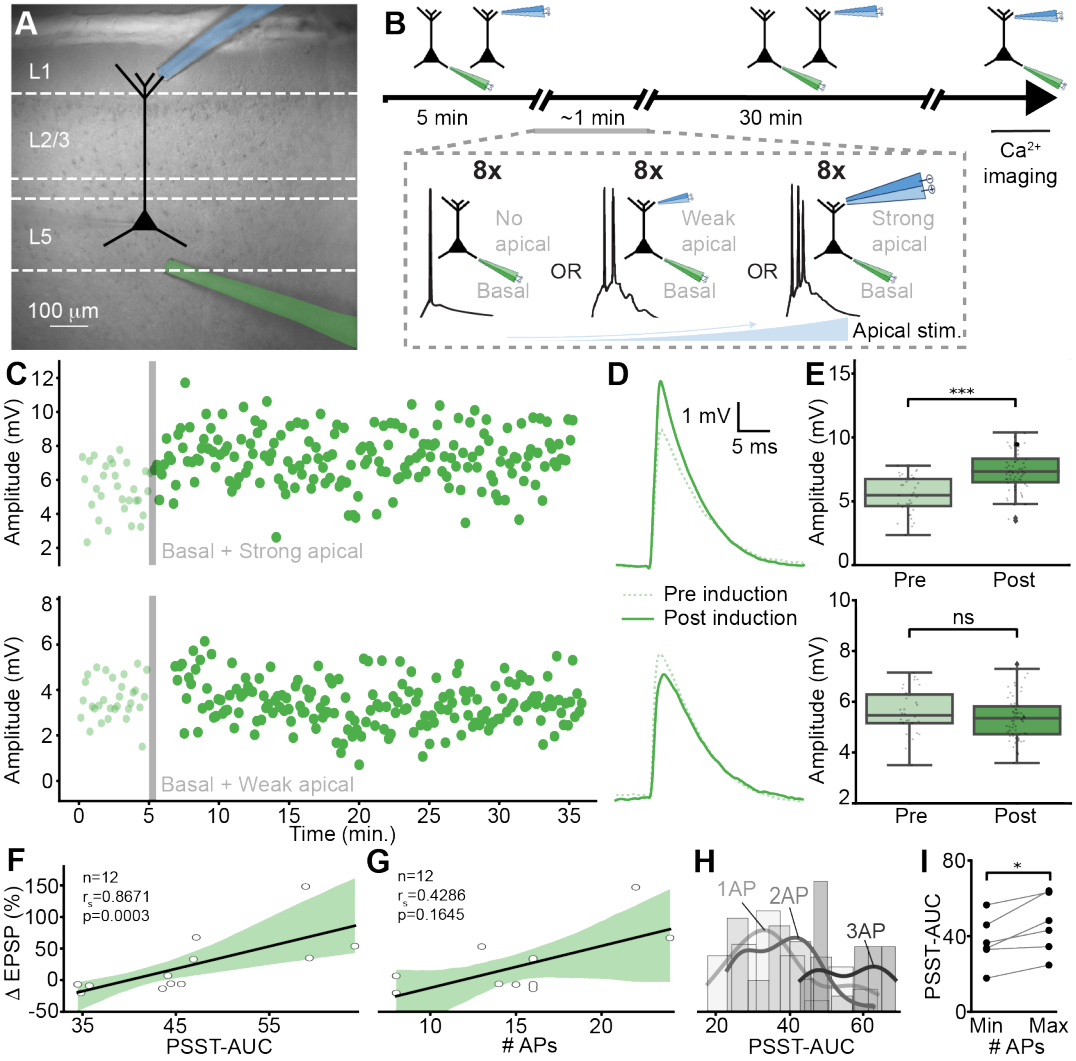

Figure 2: Apical inputs direct changes in synaptic strength in basal synapses and affect PN activity. (**A**) Wide-field image of mouse neocortex with an L5 pyramidal neuron sketch superimposed and placement of extracellular basal and apical stimulation electrodes. (**B**) Stimulation and recording protocol. (**C-D-E**) Example recordings of basal EPSP amplitudes before and after a plasticity induction event with significant increase (top) and no-significant difference in synaptic strength (bottom). (**C**) Evolution of the EPSP amplitude over time. (**D**) Mean EPSP for pre and post plasticity-induction event. (**E**) Quantification of ΔEPSP (Student T-test, top: $p < 0.0001$, bottom: $p = 0.229$), 12 neurons from 6 animals. (**F-G**) Correlation analysis between changes in basal EPSP amplitude and PSST-AUC and number of APs, respectively (Spearman's rank correlation, n=12, **F**: R_s=0.6871, p=0.0003, **G**: R_s=0.4286, p=0.1645). (**H**) Histogram of the PSST-AUC for one, two, and three APs. (**I**) Pairwise comparison of PSST-AUC for the minimum and maximum number of somatic APs (Wilcoxon signed-rank test, $p = 0.031$).

To explain the observed changes in the EPSP amplitude changes, we evaluated potential predictors of basal synaptic plasticity. We correlated the number of APs and the area under the curve of the somatic membrane depolarisation during PSSTs (PSST-AUC) to the change in EPSP amplitudes (ΔEPSP). We found a significant positive correlation between the PSST-AUC and ΔEPSP ($R_{Spearman} = 0.8671,\ p = 0.0003$, see **Fig. 2F**). In contrast, the correlation between the number of APs and ΔEPSP was $R_{Spearman} = 0.4286$ with $p = 0.1645$ (see **Fig. 2G**). This indicates that the total somatic depolarization that follows synaptic input is a better predictor of changes in synaptic strength than the number of APs following synaptic input. We note that the prolonged somatic depolarization resulting from dendritic plateau potentials can lead to burst firing (Larkum et al., 1999), indicating a dependency between the number of APs and the PSST-AUC. In this regard, we indeed find that the number of APs increases with the measured PSST-AUC (see **Fig. 2H**), and that even on a single neuron basis, the change in the number of APs significantly depends on the PSST-AUC (see **Fig. 2I**, $p = 0.031$). Our recordings also contain data for apical synapses, but in contrast to basal synaptic plasticity we find that apical synaptic strength changes do not correlate with the changes found in basal synapses (see **Supp. 5** and **Fig. S7**, $r_s = 0.2308,\ p = 0.4705$), indicating that apical synapses have different plasticity rules than basal ones.

In summary, our in vitro experiments support the predictions of our PN model by showing that basal plasticity is directed by prolonged somatic depolarization modulated by apical input. In addition, the same apical inputs induce long-lasting depolarization and trigger bursting behavior at the soma. These results agree with previous literature indicating that plasticity in PNs is reflected in their activity during learning (Larkum et al., 2004), and suggests that the implementation of BP or TL in cortical circuit models should be expressed in the PN population activity during learning.

3.3 Linking single cell synaptic plasticity to in vivo population dynamics during learning

To connect our PN model and in vitro results to in vivo population dynamics and plasticity during learning, we first recast the synaptic weight updates into a simple form of a rate-based three-factor learning rule (Frémaux and Gerstner, 2016; Gerstner et al., 2018). We then relate these factors to specific neural activity patterns that have been recorded in vivo.

In our PN model, synaptic learning is driven by the apical input as follows (see **Methods 3**),

$$\Delta w \propto r_{\text{pre}} r_{\text{post}}^{\text{baseline}} (a - a^{\text{baseline}}) \tag{6}$$

where r_{pre} is the rate of presynaptic EPSPs, $r_{\text{post}}^{\text{baseline}}$ is the postsynaptic activity at baseline apical input a^{baseline}, and a the apical input used during learning. Crucially, we model the apical input to have a multiplicative effect on the PN activity (see **Fig. 2H&I** and Larkum et al. (2004)). Thus we can rewrite the previous learning rule as,

$$\Delta w \propto r_{\text{pre}} (r_{\text{post}}^{\text{apical}} - r_{\text{post}}^{\text{baseline}}), \tag{7}$$

where $r_{\text{post}}^{\text{apical}}$ is the postsynaptic firing rate during apically-induced plasticity (see **Methods 3**). The term $r_{\text{post}}^{\text{apical}} - r_{\text{post}}^{\text{baseline}}$ can be interpreted as the subsequent change in the postsynaptic activity (see **Methods 4**). Previous works showed that the learning rule from **Eq. 7** is consistent with classical learning rules such as STDP (Aceituno et al., 2023) and with existing hierarchical cortical learning models implementing either BP or TL (Rao and Ballard, 1999; Whittington and Bogacz, 2017; Scellier and Bengio, 2017; Sacramento et al., 2018; Meulemans et al., 2021).

However, the learning rule from **Eq. 7** is similar but not identical for BP and TL. Our theoretical analysis of BP and TL algorithms reveals that the population activity during learning must differ

for both algorithms (see **Supp. 2**). For BP, the neuronal activity remains similar between inference and learning, while TL requires a target neuronal activity to be enforced during learning (see **Methods 5**). This differentiation delineates our hypotheses which can be tested with specific in vivo data. In particular, **Eq. 7** shows that the critical metric determining plasticity is the separation between sensory-driven PN activity at baseline apical input and PN activity when the apical input diverges from its baseline. Such separation can be observed in the context of neural reactivations, where neural activity seen during stimulus presentation reoccurs with slight variations when the animal is at rest or sleeping (Rothschild et al., 2017; Jadhav et al., 2012). Consistent with the feedback-driven activity of our model, cortical reactivations are required for learning (Chen and Wilson, 2023), are known to be driven by inputs to the apical dendrite (Khodagholy et al., 2017), and correlate with bursts (Sirota et al., 2003).

To test our hypotheses, we adopt data from a recent experimental study which recorded PN stimulus-evoked responses and reactivations from the mouse lateral visual cortex over learning using calcium imaging (Nguyen et al., 2024). We denote stimulus-evoked neural responses during stimulus presentations as $r_{\text{post}}^{\text{baseline}}$ and reactivations occurring during rest as $r_{\text{post}}^{\text{apical}}$ (see **Fig. 3A** and Nguyen et al. (2024) for complete details). We start by testing whether this in vivo data is consistent with our PN model. First, we verify that neuronal activity and plasticity inducing events are visible in calcium imaging of in vivo data by reproducing induction events while recording voltage and calcium signals. We found a significant correlation between somatic calcium recordings and PSST-AUC (see **Supp. 4** and **Fig. S6B**, $R_{Spearman} = 0.8303$, $p = 0.0029$). Then, we track the individual PN activities over time as the stimulus is presented, and we correlate the difference in activity between a reactivation and its associated stimulus-evoked response to the difference in activity of consecutive stimulus-evoked responses for the same stimulus. As expected (see **Supp. 1** and **Fig. S2** for all details), we find a positive correlation between these activity changes (see **Fig. 3C**, $R^2 = 0.80$), validating our PN model and **Eq. 7** for this dataset.

Thus, by leveraging **Eq. 7** and the dataset from Nguyen et al. (2024) containing both stimulus evoked responses and reactivations, we can make distinct predictions for BP and TL and derive a precise test. If the cortex implements BP, the PN activity during reactivations for a given stimulus should resemble the stimulus-evoked activity from previous stimulus presentation. If the cortex uses TL, the PN activity during reactivations should resemble the stimulus-evoked activity at the end of learning when the neural representations have stabilized.

3.4 Target learning explains plasticity in cortical networks

To test the BP and TL hypotheses on the in vivo dataset, we measure the alignment between the reactivations and the appropriate stimulus-evoked responses after learning using the cosine similarity (see **Methods 6.1&6.2**). We find that reactivations seem to align better with the stimulus-evoked responses after learning rather than the previous stimulus-evoked response (see **Fig. 3C**), as would be expected from TL (see **Supp. 2**). To evaluate which hypothesis explains the cortical reactivations better, we calculate the fraction of variance in the neuronal activity during reactivations that is explained by the BP or TL hypotheses (see **Methods 6.3**). Our results thus indicate that TL explains the activity during neural reactivations better than BP (see **Fig. 3C**, Welch T-test, $p < 0.0001$).

To ensure that our results are not influenced by the choice of baseline activity, we visualize the activity in relation to the early-to-late stimulus-evoked responses using a principal component analysis (see **Methods 6.4**). As before, we find that the reactivations are tightly concentrated around the stimulus-evoked responses after learning, as in TL. We complement this visualization by measuring the distance between the observed reactivations and the reactivations that would

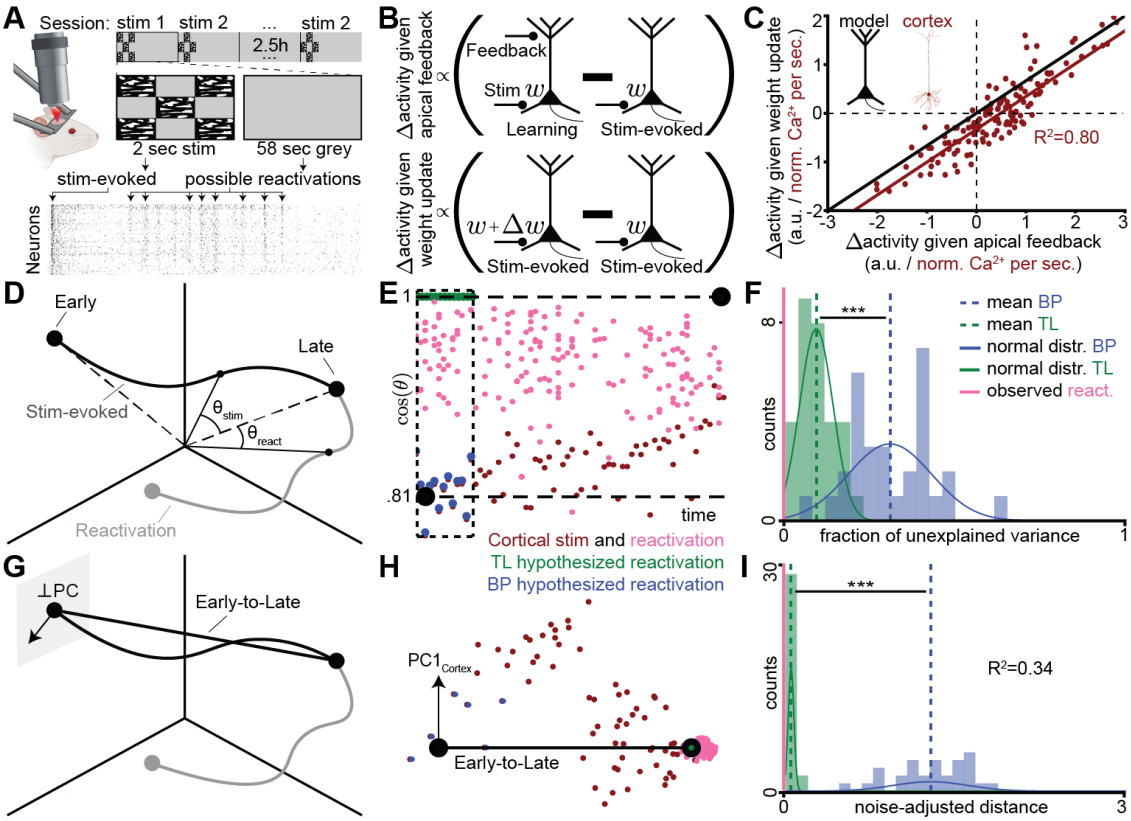

Figure 3: Target learning better aligns with in vivo cortical learning dynamics. (**A**) Experimental setup of the in vivo dataset (adapted from Nguyen et al. (2024)). Two stimuli are presented in alternation with a resting period in between. The neuronal activity is recorded through two photon population calcium imaging, and stimulus-evoked responses are used to identify possible reactivations. (**B**) Activity changes given the apical feedback are proportional to the difference between reactivations and stimulus-evoked responses (top). Activity changes given the weight update is proportional to the difference between consecutive stimulus-evoked responses (bottom). (**C**) Correlation between the changes in activity driven by apical inputs during learning against changes in activity due to synaptic updates before and after learning ($R^2 = 0.80$). (**D**) The stimulus-evoked responses and their reactivations define trajectories in activity space during learning, which we visualize through their alignment with late stimulus responses. (**E**) Cosine distance between the late stimulus-evoked responses and reactivations as well as stimulus-evoked responses. The reactivations that are hypothesized by BP and TL are shown for the first 30 minutes of recording. (**F**) The fraction of unexplained variance of the reactivations for both the BP and TL hypotheses, where the fraction is significantly lower for TL (Welch T-test, $p < 0.0001$). (**G**) Neuronal activity over time projected onto a low-dimensional space defined by the early-to-late training vector and the principal components orthogonal to this vector. (**H**) Visualization of the projected stimulus-evoked responses and PN reactivations with the hypothesized reactivations of BP and TL overlaid. (**I**) Distance between the reactivations and the TL and BP hypotheses, showing that TL predicts reactivations significantly better than BP (right, Welch T-test, $p < 0.0001$).

be expected from both hypotheses in this principal component space (see **Methods 6.5**). This analysis confirms that the reactivations are better explained by the TL hypothesis (see **Fig. 3D**, Welch T-test, $p < 0.0001$, and **Fig. S4** for the same test with more principal components).

Finally, we visualize the resemblance between cortical and artificial learning dynamics by integrating our PN model into ANNs implementing either BP and TL (see **Methods 7**). Using the same visualization as in **Fig. 3C**, we observe that the cortical learning dynamics better resemble the TL simulations rather than BP (see **Fig. S5**).

4 Discussion

Our single-cell modeling and in vitro results combined with the in vivo characterization of the PN population learning dynamics show that cortical dynamics and plasticity align more closely with target learning (TL) rather than the classical backpropagation (BP) method used in deep learning. This finding addresses a longstanding debate in neuroscience and machine learning about the type of learning algorithm the cortex employs (Grossberg, 1987; Crick, 1989; Körding and König, 2001; Richards et al., 2019; Lillicrap et al., 2020; Song et al., 2024).

In contrast to previous theoretical studies, our work focuses on formulating hypotheses on experimental data, which we verify in vitro and in vivo. Our work differs from these studies (Rao and Ballard, 1999; Scellier and Bengio, 2017; Sacramento et al., 2018; Meulemans et al., 2021, 2022b) in that we do not propose any concrete cortical circuit. Instead, we focus on comparing the dynamics of PNs and cortical networks during learning. This paradigm shift is crucial, because BP and TL are mathematical frameworks and are thus not tied to any specific formalism, network architecture or circuit. Yet, they impose different requirements on the neuronal activity. If network activity during learning is similar to the activity during stimulus presentation, then we obtain BP, whereas when activity during learning is similar to the activity after learning converges, then we define this as TL. Accordingly, many different network architectures such as control-based or predictive coding as well as energy models can implement approximations of BP or TL (Scellier and Bengio, 2017; Whittington and Bogacz, 2017; Song et al., 2024).

Importantly, our work does not characterize nor suggest any concrete circuit implementation of this family of algorithms. To address this gap, future work needs to narrow down the precise circuitry through a rigorous investigation of the feedback pathways which currently differ across TL architectures (Rao and Ballard, 1999; Scellier and Bengio, 2017; Meulemans et al., 2022b). Although we measured apical plasticity in L5 PNs directly (see **Supp. 5** and **Fig. S7**), we did not find any consistent relation between PN activity and apical plasticity, most likely due to the unspecific nature of our stimulation. As we stimulate inputs to apical dendrites using extracellular electrodes, we probably jointly activate axons from the contralateral prefrontal cortex (cPFC) as well as thalamic nuclei, specifically the mediodorsal (MD) and ventromedial (VM) thalamus. Those axonal feedback projections are likely mediated by different neuron types, which could explain the heterogeneous plasticity we observed. Future experimental research needs to disentangle the contributions of different afferent feedback projections to elicit apical input and drive plasticity, for example, using projection-specific optogenetic methods.

Cortical TL may also provide new inspiration for machine learning researchers, especially in the realm of continual learning, where target-based methods have shown great promise (Lässig et al., 2023; Song et al., 2024). However, we note that training ANNs with TL requires simulating continuous-time network dynamics by iterating feedforward and feedback passes multiple times for each data sample and is therefore computationally intensive (Rao and Ballard, 1999; Scellier and Bengio, 2017; Meulemans et al., 2021). On digital processing hardware like GPUs, this renders the

training process significantly slower and less salable compared to conventional deep learning frameworks based on BP. Nonetheless, TL research is still in its infancy and future studies might develop more efficient algorithmic implementations of TL or advance neuromorphic computing hardware to better support neuroscience-inspired algorithms.

In summary, our interdisciplinary study combines mathematical models and deep learning theory with in vitro neuroscience experiments and in vivo cortical data. Our investigation bridges multiple scales of neuronal learning suggesting that cortical hierarchical networks learn by enforcing targets rather than correcting errors, thus differing from their canonical artificial counterparts. Understanding that brains learn differently from existing machine learning models will likely inspire the development of novel and improved deep learning algorithms and hardware.

5 Author contributions

P.V.A. conceptualized the project and the modeling experiments. P.V.A. and S.d.H. developed the single neuron model and the in vivo data analysis. P.V.A. performed the mathematical analyses of the models. S.d.H. implemented and performed the simulations for the single neuron model and deep learning models, and performed the data analysis. R.L. conceptualized and performed the electrophysiological experiments. R.L. provided biology insights about the project on both theoretical and experimental neuroscience. B.F.G. supervised the project. P.V.A., S.d.H., R.L., and B.F.G. wrote the paper.

6 Author information

The authors have declared no competing interest.

7 Data and code availability statement

Experimental data and analysis code will be made public upon acceptance

8 Acknowledgements

This work was supported by the Swiss National Science Foundation (B.F.G. CRSII5-173721 and 315230 189251, P.V.A. 182539), ETH project funding (B.F.G. ETH-20 19-01), and a Forschungskredit from the University of Zürich (P.V.A).

References

Aceituno, P. V., Farinha, M. T., Loidl, R., and Grewe, B. F. (2023). Learning Cortical Hierarchies with Temporal Hebbian Updates. *bioRxiv* doi:10.1101/2023.01.02.522459

Artola, A. and Singer, W. (1987). Long-term potentiation and nmda receptors in rat visual cortex. *Nature* 330, 649–652

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711

Bear, M. F., Press, W. A., and Connors, B. W. (1992). Long-term potentiation in slices of kitten visual cortex and the effects of nmda receptor blockade. *Journal of Neurophysiology* 67, 841–851

Bengio, Y., Lee, D.-H., Bornschein, J., Mesnard, T., and Lin, Z. (2016). Towards Biologically Plausible Deep Learning. *arXiv* doi:10.48550/arXiv.1502.04156

Chen, Z. S. and Wilson, M. A. (2023). How our understanding of memory replay evolves. *Journal of Neurophysiology* 129, 552–580

Chindemi, G., Abdellah, M., Amsalem, O., Benavides-Piccione, R., Delattre, V., Doron, M., et al. (2022). A calcium-based plasticity model for predicting long-term potentiation and depression in the neocortex. *Nature Communications* 13, 3038. doi:10.1038/s41467-022-30214-w

Coogan, T. A. and Burkhalter, A. (2004). Conserved patterns of cortico-cortical connections define areal hierarchy in rat visual cortex. *Experimental Brain Research* 80, 49–53

Crick, F. (1989). The recent excitement about neural networks. *Nature* 337, 129–132. doi: 10.1038/337129a0

D'Souza, R. D. and Burkhalter, A. (2017). A laminar organization for selective cortico-cortical communication. *Frontiers in Neuroanatomy* 11

Emptage, N., Bliss, T. V., and Fine, A. (1999). Single synaptic events evoke nmda receptor–mediated release of calcium from internal stores in hippocampal dendritic spines. *Neuron* 22, 115–124

Evans, R. and Blackwell, K. (2015). Calcium: Amplitude, Duration, or Location? *The Biological bulletin* 228, 75–83

Frémaux, N. and Gerstner, W. (2016). Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Frontiers in neural circuits* 9, 85

Gerstner, W., Lehmann, M., Liakoni, V., Corneil, D., and Brea, J. (2018). Eligibility traces and plasticity on behavioral time scales: experimental support of neohebbian three-factor learning rules. *Frontiers in neural circuits* 12, 53

Graupner, M. and Brunel, N. (2012). Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proceedings of the National Academy of Sciences* 109, 3991–3996. doi:10.1073/pnas.1109359109

Grewe, B., Bonnan, A., and Frick, A. (2010). Back-propagation of physiological action potential output in dendrites of slender-tufted L5A pyramidal neurons. *Frontiers in Cellular Neuroscience* 4

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

Gulledge, A. T. and Stuart, G. J. (2003). Action Potential Initiation and Propagation in Layer 5 Pyramidal Neurons of the Rat Prefrontal Cortex: Absence of Dopamine Modulation. *The Journal of Neuroscience* 23, 11363–11372. doi:10.1523/JNEUROSCI.23-36-11363.2003

Inglebert, Y., Aljadeff, J., Brunel, N., and Debanne, D. (2020). Synaptic plasticity rules with physiological calcium levels. *Proceedings of the National Academy of Sciences of the United States of America* 117, 33639–33648. doi:10.1073/pnas.2013663117

Jadhav, S. P., Kemere, C., German, P. W., and Frank, L. M. (2012). Awake hippocampal sharp-wave ripples support spatial memory. *Science* 336, 1454–1458

Kasper, E. M., Larkman, A. U., Lübke, J., and Blakemore, C. (1994). Pyramidal neurons in layer 5 of the rat visual cortex. I. Correlation among cell morphology, intrinsic electrophysiological properties, and axon targets. *The Journal of Comparative Neurology* 339, 459–474. doi: 10.1002/cne.903390402

Khaligh-Razavi, S.-M. and Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain it cortical representation. *PLoS computational biology* 10, e1003915

Khodagholy, D., Gelinas, J. N., and Buzsáki, G. (2017). Learning-enhanced coupling between ripple oscillations in association cortices and hippocampus. *Science* 358, 369–372

Körding, K. P. and König, P. (2001). Supervised and Unsupervised Learning with Two Sites of Synaptic Integration. *Journal of Computational Neuroscience* 11, 207–215. doi: 10.1023/A:1013776130161

Larkum, M., Zhu, J., and B., S. (2001). Dendritic mechanisms underlying the coupling of the dendritic with the axonal action potential initiation zone of adult rat layer 5 pyramidal neurons. *The Journal of physiology* 533, 447–66. doi:10.1111/j.1469-7793.2001.0447a.x

Larkum, M. E., Senn, W., and Lüscher, H.-R. (2004). Top-down dendritic input increases the gain of layer 5 pyramidal neurons. *Cerebral Cortex (New York, N.Y.: 1991)* 14, 1059–1070. doi:10.1093/cercor/bhh065

Larkum, M. E., Zhu, J. J., and Sakmann, B. (1999). A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature* 398, 338–341. doi:10.1038/18686

Lässig, F., Aceituno, P. V., Sorbaro, M., and Grewe, B. F. (2023). Bio-inspired, task-free continual learning through activity regularization. *Biological Cybernetics* 117, 345–361

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. doi: 10.1038/ncomms13276

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., and Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience* 21, 335–346

Lisman, J. (1989). A mechanism for the Hebb and the anti-Hebb processes underlying learning and memory. *Proceedings of the National Academy of Sciences of the United States of America* 86, 9574–9578. doi:10.1073/pnas.86.23.9574

Majewska, A., Brown, E., Ross, J., and Yuste, R. (2000). Mechanisms of Calcium Decay Kinetics in Hippocampal Spines: Role of Spine Calcium Pumps and Calcium Diffusion through the Spine Neck in Biochemical Compartmentalization. *The Journal of Neuroscience* 20, 1722–1734. doi: 10.1523/JNEUROSCI.20-05-01722.2000

Markram, H. (1997). A network of tufted layer 5 pyramidal neurons. *Cerebral cortex (New York, NY: 1991)* 7, 523–533

Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science (New York, N.Y.)* 275, 213–215. doi: 10.1126/science.275.5297.213

McRory, J. E., Santi, C. M., Hamming, K. S. C., Mezeyova, J., Sutton, K. G., Baillie, D. L., et al. (2001). Molecular and Functional Characterization of a Family of Rat Brain T-type Calcium Channels. *Journal of Biological Chemistry* 276, 3999–4011. doi:10.1074/jbc.M008215200

Meulemans, A., Farinha, M. T., Cervera, M. R., Sacramento, J., and Grewe, B. F. (2022a). Minimizing control for credit assignment with strong feedback. In *International Conference on Machine Learning* (PMLR), 15458–15483

Meulemans, A., Tristany Farinha, M., Garcia Ordonez, J., Vilimelis Aceituno, P., Sacramento, J. a., and Grewe, B. F. (2021). Credit assignment in neural networks through deep feedback control. In *Advances in Neural Information Processing Systems* (Curran Associates, Inc.), vol. 34, 4674–4687

Meulemans, A., Zucchet, N., Kobayashi, S., Von Oswald, J., and Sacramento, J. (2022b). The least-control principle for local learning at equilibrium. *Advances in Neural Information Processing Systems* 35, 33603–33617

Nevian, T. and Sakmann, B. (2006). Spine Ca2+ Signaling in Spike-Timing-Dependent Plasticity. *The Journal of Neuroscience* 26, 11001–11013. doi:10.1523/JNEUROSCI.1749-06.2006

Nguyen, N. D., Lutas, A., Amsalem, O., Fernando, J., Ahn, A. Y.-E., Hakim, R., et al. (2024). Cortical reactivations predict future sensory responses. *Nature* 625, 110–118

Payeur, A., Guerguiev, J., Zenke, F., Richards, B. A., and Naud, R. (2021). Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. *Nature Neuroscience* 24, 1010–1019. doi:10.1038/s41593-021-00857-x

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, 79–87. doi:10.1038/4580

Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., et al. (2019). A deep learning framework for neuroscience. *Nature neuroscience* 22, 1761–1770

Rothschild, G., Eban, E., and Frank, L. M. (2017). A cortical–hippocampal–cortical loop of information processing during memory consolidation. *Nature neuroscience* 20, 251–259

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

Saxe, A., Nelli, S., and Summerfield, C. (2021). If deep learning is the answer, what is the question? *Nature Reviews Neuroscience* 22, 55–67

Scellier, B. and Bengio, Y. (2017). Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in computational neuroscience* 11, 24

Sirota, A., Csicsvari, J., Buhl, D., and Buzsáki, G. (2003). Communication between neocortex and hippocampus during sleep in rodents. *Proceedings of the National Academy of Sciences* 100, 2065–2069

Sjöström, J., Gerstner, W., et al. (2010). Spike-timing dependent plasticity. *Spike-timing dependent plasticity* 35, 0–0

Sjöström, P. J., Turrigiano, G. G., and Nelson, S. B. (2003). Neocortical ltd via coincident activation of presynaptic nmda and cannabinoid receptors. *Neuron* 39, 641–654

Slutsky, I., Sadeghpour, S., Li, B., and Liu, G. (2004). Enhancement of Synaptic Plasticity through Chronically Reduced Ca2+ Flux during Uncorrelated Activity. *Neuron* 44, 835–849. doi:10.1016/j.neuron.2004.11.013

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* , 1–11

Stuart, G., Schiller, J., and Sakmann, B. (1997). Action potential initiation and propagation in rat neocortical pyramidal neurons. *The Journal of Physiology* 505, 617–632

Stuart, G. J. and Sakmann, B. (1994). Active propagation of somatic action potentials into neocortical pyramidal cell dendrites. *Nature* 367, 69–72. doi:10.1038/367069a0

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, 1229–1262. doi:10.1162/NECO_a_00949

Wimmer, V., Bruno, R., Kock, C., Kuner, T., and Sakmann, B. (2010). Dimensions of a projection column and architecture of vpm and pom axons in rat vibrissal cortex. *Cerebral cortex (New York, N.Y. : 1991)* 20, 2265–76. doi:10.1093/cercor/bhq068

Yang, S. N., Tang, Y. G., and Zucker, R. S. (1999). Selective induction of LTP and LTD by postsynaptic [Ca2+]i elevation. *Journal of Neurophysiology* 81, 781–787. doi:10.1152/jn.1999.81.2.781

Supplementary Materials for

Challenging Backpropagation: Evidence for Target Learning in the Cortex

This PDF file includes:

Methods	2
Supplementary Text	10
Figs. S1 to S7	18

Methods

1 Units and model parameters

The parameters in our model are collected in Table 1.

2 Electrophysiology

Male B6/J-Crl1 mice of strain C57BL/6J (orig. Charles River) were bred and maintained at our in-house facility. Animals were kept in their home cages under normal day/night conditions (light: 7am – 7pm). Mouse experiments were approved by the institutional animal care and use committee of Anonymous Institution. All experiments were conducted with the approval of the Anonymous State Veterinary Office.

Mice (11-13 weeks old) were deeply anesthetized using Isoflurane and decapitated. The brains were dissected from the skull and submerged in ice-cold high-sucrose artificial cerebrospinal fluid (ACSF) during acute slice preparation. High sucrose ACSF containing (in mM): 75 Sucrose, 10 Glucose, 80 NaCl, 26 NaHCO3, 7 MgSO4, 2.5 KCl, 1 NaH2PO4, 0.5 CaCl2, fumigated with oxycarbon (95% O2, 5% CO2). Coronal slices (300 µm) were drawn from PFC (+1.4 to +2.6 from bregma) using a VT1200 S vibratome (Leica) and incubated for 30 minutes in oxygenated holding ACSF at 36°C. Holding ACSF containing (in mM): 15 Glucose, 120 NaCl, 26 NaHCO3, 0.9 MgSO4, 2.5 KCl, 1.25 NaH2PO4, 1.3 CaCl2, fumigated with oxycarbon (95% O2, 5% CO2). Slices rested for at least 45 minutes in holding ACSF at room temperature prior to electrophysiological recordings.

Neurons were visualized with differential interference contrast (DIC) using an Olympus BX61 WI microscope with an Olympus 60x water immersion objective and a CellCam Kikker MT100 (Cairn) DCIM camera. L5 pyramidal neurons (PN) in the PFC were identified based on morphology and distance to the pia and patched with glass pipettes with resistances of 4.8-6.4 MΩ filled with intracellular solution containing (in mM): 10 HEPES, 20 KCl, 117 K-gluconate, 4 Mg-ATP, 0.3 GTP, 10 Na-P-creatine. For morphology reconstruction, the intracellular solution also contained 0.2% Biocytin. For Ca^{2+} imaging experiments, the intracellular solution also contained 200 µM Oregon Green™ 488 BAPTA-5N. Whole-cell, somatic patch-clamp recordings were acquired in current-clamp mode using an Axon MultiClamp 700A amplifier and digitized using an Axon DigiData 1550 B digital converter. Data was sampled at 50 kHz and filtered at 3 kHz. Typical resting membrane potentials were between -64mV and -72mV, and access resistance ranged from 15 MΩ to 25 MΩ. Recordings in which neurons depolarized > 10% from resting potential or access resistance changed > 15% from initial value were discarded. Bridge potentials were compensated, and liquid-junction potentials were not corrected for. Depolarizing and hyperpolarizing currents were injected to characterize neurons further. Only neurons showing L5 PN-typical I_h-currents were considered. Theta-barrel glass-pipettes with openings of 30 µm were placed at 50 µm – 80 µm from the patched neuron to stimulate afferents targeting basal dendrites as well as in cortical layer 1 to stimulate apical afferents with synapses onto apical tuft dendrites.

After stabilization of resting membrane potential, apical and basal stimulation strength was adjusted to elicit EPSPs with amplitudes in the range of 2-3 mV. This usually required 100 µs voltage pulses in the range of -1 to -3 V followed by immediate 100 µs voltage pulses of +1 to +3 V for basal stimulations. Apical stimulation had the same shape but ranged from -6 to -8 V and +6 to +8 V, respectively. Next, the required strength of basal stimulations to induce a single action potential was determined. For this, tripling basal stimulation amplitudes and stimulating 3 times at 100 Hz was usually sufficient. If extracellular electrical stimulation was not sufficient to trigger action potentials, additional current was injected into the soma via the patch pipette until an AP was induced. Finally, apical stimulation strength was increased until multiple action potentials or plateau potentials were induced.

After all required stimulation amplitudes were set, baseline EPSP amplitudes were recorded for 5 minutes with alternating basal and apical stimulations every 5 seconds followed by 8 combined, strong basal and apical stimulations with a 10 second interval, each eliciting a plasticity induction event. Next, EPSPs were recorded for at least 30 minutes (same as baseline recordings).

Analysis of electrophysiology traces was performed using individual Python scripts and the pyabf library. The correlation analyses were performed with the seaborn library (regplot) and the statistical analysis with the scipy library (spearmanr).

Analysis of PSST-AUC with different numbers of action potentials (**Fig. 2K**) was performed on a subset of neurons in which PSSTs occurred with different amounts of APs (1vs2, 1vs3, or 2vs3). PSST-AUCs were averaged for events with min. and max. number of APs respectively and plotted as 2 data points per cell.

In the case of cell morphology reconstruction (only performed in 8 of 12 neurons), slices were kept in 4% PFA for 24-32 hours and stained overnight with Streptavidin, Alexa Fluor 350 conjugate (invitrogen) and mounted on commercially available microscopy slides.

Calcium imaging was performed in a subset of neurons (n = 7 neurons, 4 mice) after post-induction EPSP recordings where completed by imaging pipette-filled OGB5-N using a Lumencor SPECTRA Light Engine (GFP/FITC 475/28 nm) for excitation and a ORCA-Lightning Digital CMOS camera (C14120-20P, Hamamatsu) for image acquisition at 50 Hz. Imaging data was pre-processed (background-subtraction, bleaching correction) using Fiji/ImageJ2. Extracted fluorescence traces were further processed and analysed using individual Python scripts. A correlation analysis between somatic calcium recordings (derived as ΔF/F AUC) and somatic membrane potential AUC showed a significant correlation ($R_{Spearman} = 0.8303$, $p = 0.0029$) and was used to infer calcium levels during the actual plasticity induction events using linear regression.

Group data are expressed as mean ± s.e.m. unless otherwise stated. Statistical significance was calculated using paired or unpaired t-tests as well as Spearman correlation.

3 Derivation of the rate-based neuron model

To render our neuron model comparable to the average activities observed in neuronal recordings in vivo and to implement the model in artificial neural networks, we reformulate our PN model as rate-based with a three-factor learning rule.

The inputs that arrive at the soma and the inputs that arrive at the apical dendrites are assumed to follow a Poisson distribution with a relatively low firing rate, in order to ensure that consecutive presynaptic spikes or feedback spikes are temporally separated with high probability, allowing us to consider spikes and bursts as independent events.

We divide the somatic output into bursts and single spikes. The number of events where somatic presynaptic input drove the soma over the threshold θ_{soma} is represented by N. The events will sometimes generate a burst N_{burst} and sometimes a single spike N_{spike}, thus $N = N_{\text{burst}} + N_{\text{spike}}$. A burst is generated only if feedback input arrives at the apical dendrites within a short time interval ΔT after a somatic postsynaptic spike. Using the aforementioned assumption of low firing rates for apical and somatic inputs, which implies that somatic spikes cannot arrive very close in time, coincidence of basal and apical inputs are independent events. We can thus calculate the expected number of bursts and single APs as,

$$N_{\text{burst}} = \frac{\Delta T}{T} aN, \quad N_{\text{spike}} = (1 - \frac{\Delta T}{T} a)N \tag{8}$$

where T represents the total time of the simulation, and a is the rate of feedback input. Intuitively, the number of bursts is simply the fraction of time at which the apical input can generate a burst times the probability of receiving suprathreshold somatic activity.

Lastly, the firing rate of the postsynaptic neuron increases when burst firing occurs. Therefore, considering the firing rate as the number of spikes, the firing rate can be expressed as,

$$r_{\text{post}} = N_{\text{spike}} + \overline{m} N_{\text{burst}} = N\left(\overline{m} + \frac{\Delta T}{T} a(1-\overline{m})\right). \tag{9}$$

where $\overline{m} > 1$ is the expected number of spikes per burst. As denoted here, the expected firing rate of the neuron grows monotonically with the amount of apical feedback.

We can now derive plasticity from the calcium model, understanding that a plateau potential generates LTP, and a burst and a single spike without the corresponding apical input generates LTD when there is a presynaptic spike

$$\Delta w = \left[\Delta_{\text{burst}} w N_{\text{burst}} - \Delta_{\text{spike}} w N_{\text{spike}}\right] \frac{N_{\text{pre}}}{T} \tag{10}$$

where $\Delta_{\text{burst}} w > 0$ and $\Delta_{\text{spike}} w > 0$ denote the synaptic change per burst or spike, respectively, and N_{pre} correspond to the number of presynaptic spikes. Substituting **Eq. 8** into **Eq. 10** produces,

$$\Delta w = \left(\frac{\Delta T}{T} N \frac{N_{\text{pre}}}{T}\right)\left(a(\Delta_{\text{burst}} w + \Delta_{\text{spike}} w) - \frac{T}{\Delta T}\Delta_{\text{spike}} w\right). \tag{11}$$

Now there exist a a^{baseline} such that $\Delta w = 0$, which corresponds to

$$a^{\text{baseline}} = \frac{\frac{T}{\Delta T}\Delta_{\text{spike}} w}{\Delta_{\text{burst}} w + \Delta_{\text{spike}} w}. \tag{12}$$

Using a^{baseline}, we can write the synaptic weight change as

$$\Delta w = \eta \frac{N_{\text{pre}}}{T} \frac{N}{T}(a - a^{\text{baseline}}), \tag{13}$$

where $\eta = \Delta T(\Delta_{\text{burst}} w - \Delta_{\text{spike}} w)$. To simplify the notation further, we will denote the presynaptic firing rate as $r_{\text{pre}} = N_{\text{pre}}/T$ and the postsynaptic firing rate at baseline apical input $r_{\text{post}}^{\text{baseline}}$, which leaves us with the synaptic update rule

$$\Delta w = \eta r_{\text{pre}} r_{\text{post}}^{\text{baseline}}(a - a^{\text{baseline}}). \tag{14}$$

and if we set $a^{\text{baseline}} = 1$ for simplicity, the postsynaptic firing rate for a given apical input is then

$$r_{\text{post}} = r_{\text{post}}^{\text{baseline}} a. \tag{15}$$

To visualize the connection between the firing rates, synaptic plasticity and apical input, we simulated our spike-based pyramidal neuron model with basal and apical inputs in **Fig. S3**.

4 From synaptic plasticity rules to learning rules

Our neuron model can be used in a deep learning setting, as long as the appropriate learning signals are provided. Note that we are not concerned with *how* the signals are computed, we simply assume that a neuronal circuit exists which delivers the signal to the apical dendrite.

The learning algorithms that we consider use synaptic update rules with a generalized Hebbian-like structure of the form

$$\Delta_{\text{alg}} w = -\eta r_{\text{pre}} \delta_{\text{post}} \tag{16}$$

where $\delta_{\text{post}} = r_{\text{post}}^{\text{apical}} - r_{\text{post}}^{\text{baseline}}$ is a value computed at the postsynaptic neuron. In the BP algorithm, this would be the derivative of the loss function with respect to the postsynaptic activity, and in target-based methods the difference between the target activity and the current activity (Meulemans et al., 2021; Rao and Ballard, 1999; Song et al., 2024), or the temporal change of the activity when driven towards the target (Aceituno et al., 2023).

We can now consider the plasticity rate of our model taken from **Eq. 14** and compute the required algorithmic apical input as

$$a_{\text{alg}} = -\eta_{Ca^{++}} \frac{\delta_{\text{post}}}{r_{\text{post}}^{\text{baseline}}} + a^{\text{baseline}} \tag{17}$$

where a^{baseline} is the apical input baseline and $r_{\text{post}}^{\text{baseline}}$ the firing rate at baseline apical input, and $\eta_{Ca^{++}}$ is the learning rate, which would either be an hyperparameter in the training process or could be extracted from our data.

Note that our equation has a singularity for $r_{\text{post}} = 0$, where the fraction would go to infinity. However, this is solved by noting that when $r_{\text{post}} = 0$ there should be no plasticity, neither from the biological point of view nor from the algorithmic one. From the biological point of view, this would naturally lead to a lack of plasticity, because there is no weight update in the absence of postsynaptic firing. From the algorithmic point of view, the use rectified linear units (or similar semi-positive functions) implies that in those cases the value of δ_{post} would be zero because a small change in the inputs to the neuron would lead to no change at all in the output. Therefore, calculus-based methods would give this neuron a derivative of zero (and hence $\delta_{\text{post}} = 0$).

5 Learning signals and reactivations for BP and TL

In this section we show how BP and TL change the dynamics during learning, and derive hypotheses for how the activity in both algorithms changes due to apical feedback.

When we inject feedback into our PNs, we automatically change their activity and enforce a change in the output of the network. We can calculate this change easily with the baseline $a^{\text{baseline}} = 1$,

$$\mathbf{r}^{\text{apical}} = \mathbf{r}^{\text{baseline}} \odot \mathbf{a} = -\mathbf{r}^{\text{baseline}} \odot \delta \oslash \mathbf{r}^{\text{baseline}} + \mathbf{r}^{\text{baseline}} = \mathbf{r}^{\text{baseline}} - \eta_{Ca^{++}}\delta \tag{18}$$

where $\mathbf{r}^{\text{apical}}, \mathbf{r}^{\text{baseline}}$ stand for the firing rates at of the neural population, , $\mathbf{a}$ for the apical input at each neuron during learning, $\odot$ and $\oslash$ are the elementwise multiplication and division respectively and δ is a term that represents how each neuron should change its activity that appears in both the BP and TL formalism (Rao and Ballard, 1999; Meulemans et al., 2021; Richards et al., 2019; Bengio et al., 2016).

5.1 BP with our PN model

The modified activity at layer alters the output of the network, reducing the error at the output layer $\mathbf{r}_L$. It does so by considering the stimulus-evoked activity and evaluating how it should change,

$$\mathbf{r}_L \to \mathbf{r}_L - \frac{\partial \mathbf{r}_L}{\partial \mathbf{r}} \delta \tag{19}$$

In BP,

$$\delta_{BP} = \left(\frac{\partial \mathbf{r}_L}{\partial \mathbf{r}} \right)^{\top} \mathbf{e}. \tag{20}$$

where the partial derivative is evaluated at the neuronal activity at inference. If we now put this into our PN model as a learning signal, the postsynaptic activity becomes

$$\mathbf{r}_L \xrightarrow[\text{BP}]{} \mathbf{r}_L - \eta_{Ca^{++}} \frac{\partial \mathbf{r}_L}{\partial \mathbf{r}} \left(\frac{\partial \mathbf{r}_L}{\partial \mathbf{r}} \right)^{\top} \mathbf{e}, \tag{21}$$

where $\eta_{Ca^{++}}$ is a scaling parameter.

Activity during learning: The synaptic update from BP is taken as the derivative of the loss with respect to the synaptic weights assuming that the neuronal activity is the one given at inference. Hence, presynaptic activity during learning is the same as in the feedforward pass, and thus the presynaptic activity during learning must be similar to the presynaptic activity during inference. As almost all the neurons in the network are presynaptic neurons for some synapse going into some other neuron, the scaling of the changes quantified by $\eta_{Ca^{++}}$ must be small.

5.2 TL with our neuron model

In TL enforce a target we get the target value first and then recover the δ. This can be formalized as

$$\mathbf{r}_L \xrightarrow[\text{TL}]{} \mathbf{r}_L - \mathbf{e} = \mathbf{r}_L - \frac{\partial \mathbf{r}_L}{\partial \mathbf{r}} \delta_{TL} \tag{22}$$

$$\delta_{TL} \approx \left(\frac{\partial \mathbf{r}_L}{\partial \mathbf{r}} \right)^{-1} \mathbf{e} \tag{23}$$

where the superscript $^{-1}$ indicates an inverse. Note that the error in here is a mathematical construct which does not need to have any explicit representation. The network would only have a response to a stimulus and a target, but the difference between them need not to be computed by any neuron.

Activity during learning: In TL, the term δ is not computed directly, but emerges from enforcing the target activity. This naturally gives a strict hypothesis for TL, namely that the activity during learning must be precisely the target activity. Since the target activity is observed after learning is complete, our hypothesis for TL is that the activity during learning resembles the activity during inference after learning.

6 In vivo population data analysis

6.1 Data source and preprocessing

We measure the alignment between the activity during learning and during inference or stimulus presentation in the cortical in vivo data available from Nguyen et al. (2024). We modify the existing code to fit the analysis described in the main text and keep the pre-processing as in the original publication (Nguyen et al., 2024). For all cortical data included in our analysis, we used only the top ten percent of neurons responding to the given stimulus, as in Nguyen et al. (2024). We denote early responses as the average neural activity of the first three stimulus-evoked responses, and late responses as the average neural activity of the last three stimulus-evoked responses. We note that neural the activity has a significant variance, and thus the three points that define the centroid have an angle of 24 degrees with the centroid itself (see the three rightmost responses in **Fig. 3B**). We will use $\mathbf{r}$ to denote the stimulus-evoked responses and ρ to denote reactivations.

6.2 Measuring alignment

To compare the reactivations with the stimulus-evoked responses we consider them as vectors in a space where every dimension corresponds to the activity of one neuron. In this vector space, we can define both the reactivations and the activities, but the baseline activity for single cells during rest periods – which is when the reactivations take place – is not the same as during active periods (Nguyen et al., 2024).

To avoid the effects of the different baselines, we measure the alignment between reactivations and late responses using the cosine distance between a given reactivation ρ_i and the reactivation

that would be given by the hypothesis ρ_i^{hyp} for each sample i,

$$\cos(\theta_i^{\text{hyp}}) = \frac{\langle \rho_i, \rho_i^{\text{hyp}} \rangle}{\|\rho_i\| \|\rho_i^{\text{hyp}}\|}, \tag{24}$$

$\langle \cdot, \cdot \rangle$ for the dot product and $\|\|$ the Euclidean norm.

Now we need to formulate our hypotheses in terms of ρ_i^{hyp}. TL would predict that the reactivations are aligned with the stimulus-evoked responses after learning, when the representations have converges, thus

$$\rho_i^{\text{TL}} \| \mathbf{r}_{\text{late}}, \tag{25}$$

where $\mathbf{r}_{\text{late}}$ refers to the reactivation after learning is finished, which we estimate by taking the average of the last three stimulus-evoked responses, and $\|$ stands for colinearity (because the baseline activities are different). Note that for TL the ρ_i^{TL} is the same for every sample. In contrast, BP would predict that the reactivations are aligned with the stimulus-evoked responses of the last stimulus presentation

$$\rho_i^{\text{BP}} \| \mathbf{r}_i. \tag{26}$$

where $\mathbf{r}_i$ refers to the stimulus evoked response directly preceding $\rho(i)$. The measure of alignment with TL is shown in **Fig. 3C** using data from sample mouse 'NN8' for stimulus 2 from the recording date 12-03-2021. As reported in the original publication (Nguyen et al., 2024), the remaining data look similar.

6.3 Hypothesis comparison through alignment

To compare the TL and BP hypotheses we focus on the variance explained by each hypothesis. We start by formulating the reactivations as neural activity under a given hypothesis as

$$\rho_i = \rho_i^{\text{hyp}} + \varepsilon_i, \tag{27}$$

where ε_i stands for the noise. Notice that since we do not have a consistent baseline for reactivations and must use the alignment or fit a scaling term, we implicitly assume that $\varepsilon_i \perp \rho_i^{\text{hyp}}$. In other words, we are masking any noise that is co-linear with our hypothesis. Note though that if we have a high dimensional neural space, the fraction of noise that would align with our hypothesis should be very small, because in high dimensional spaces a random vector (ε_i) would be almost orthogonal to a specific direction (ρ_i^{hyp}).

We can thus measure the fraction of variance unexplained in each reactivation that can be obtained as

$$\frac{\|\varepsilon_i\|^2}{\|\rho_i\|^2} = \sin^2(\theta_i^{\text{hyp}}) = 1 - \cos^2(\theta_i^{\text{hyp}}) = 1 - \left(\frac{\langle \rho_i, \rho_i^{\text{hyp}} \rangle}{\|\rho_i\| \|\rho_i^{\text{hyp}}\|} \right)^2. \tag{28}$$

We thus obtain a value for each reactivation. We perform a Welch t-test to assess the significance of the difference.

6.4 Origin-independent analysis

A potential concern with using the cosine similarity is that the results might change depending on the choice of origin of the neuronal activity. To address this concern, we also evaluate the variance across the stimulus-evoked responses and reactivations by projecting the variance that is not explained by the change of early-to-late activity. Since the relative position of the early and late activities is independent of the choice of basis or origin, this ensures that our results are consistent even if the origin was chosen differently.

We start by computing the difference vector between early and late responses as,

$$\mathbf{d} = \mathbf{r}_{\text{late}} - \mathbf{r}_{\text{early}}, \tag{29}$$

and then evaluate the variance of the activity that is not explained by this difference. We do this by projecting the activity on the subspace orthogonal to the early-to-late vector, and then we compute the first principal component (PC) and project the neural activity for both reactivations and stimulus-evoked activity onto the component

$$\mathbf{r}_{\perp\mathbf{d}} = \mathbf{r} - \mathbf{d}\frac{\langle \mathbf{r}, \mathbf{d}\rangle}{\|\mathbf{d}\|^2}, \quad \rho_{\perp\mathbf{d}} = s\rho - \mathbf{d}\frac{\langle s\rho, \mathbf{d}\rangle}{\|\mathbf{d}\|^2}. \tag{30}$$

Note that since the reactivations have a different baseline, we added a constant scaling term $s = 1.3$ that ensures that the reactivations have the same magnitude as the stimulus-evoked activities, which we take from the original dataset analysis (Nguyen et al., 2024).

6.5 Hypothesis comparison through distances

We can use projection onto the early-to-late vector and PC components orthogonal to it as a dimensionality reduction procedure that maintains the space where our hypotheses live (the early-to-late vector) as well as other neural variability. In this space, we can measure the distance between the activity that would be expected in both our hypotheses and what we observe in the data.

Our approach is thus to project the neural activity of the reactivations into the space given by

$$\vec{u}_{\mathbf{d}} = \frac{\mathbf{d}}{\|\mathbf{d}\|}, \ \vec{u}_{\perp\mathbf{d}} = PC(\rho_{\perp\mathbf{d}}) \tag{31}$$

where PC is the principal component operator that is applied to the data projected on the component orthogonal do $\mathbf{d}$. In this new basis, we can measure the noise-adjusted Euclidean distances between the points

$$p^{\text{BP}} = (0, 0), \quad p^{\text{TL}} = (\|\mathbf{d}\|, 0), \quad D_{\text{SNR}}(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} \frac{(x_i - y_i)^2}{\sigma_i^2}} \tag{32}$$

where each one corresponds to the BP or TL hypotheses. We can thus obtain a distance value for each reactivation for each hypotheses, and we compare the distributions generated by both through a Welch t-test.

7 Network simulations

We constructed ANNs with three hidden layers containing 100 hidden neurons each with ReLU activation functions. The hidden neurons were modified to be our rate-based multiplicative single neuron model (see **Methods 3**). Where the teaching signals for BP and TL are landing on the apical dendrite, and for BP modified according to **Methods 4**.

We selected a supervised classification task for training for two main reasons. First, previous studies have noted that supervised learning for categories leads to feature representations in neural activity that are similar to those found in cortex, whereas unsupervised methods do not (Khaligh-Razavi and Kriegeskorte, 2014). Second, the cortical recordings that we use do suggest that cortex is learning a separation of representations of two distinct stimuli (Nguyen et al., 2024) which is directly transferable to a supervised classification task, whereas for unsupervised training we would have to use a loss function that we do not know for cortex.

The TL experiments shown in the main text were conducted through the Deep Feedback Control (DFC) framework with strong feedback (Meulemans et al., 2021, 2022a). This architecture was

chosen because DFC does not require a symmetric or anti-symmetric weight structure as other methods Whittington and Bogacz (2017); Scellier and Bengio (2017), which is hard to implement when the feedforward inputs are additive but the feedback is multiplicative. For the BP experiments we used the standard PyTorch implementation of BP, extracted the learning signal for every neuron and sent the signal to each neuron.

For training, we aimed to follow Nguyen et al. (2024) as closely as possible. Considering a binary stimulus task, where the network is presented with two of the MNIST digits in an alternating fashion for 80 stimuli with a learning rate of $\eta = 0.001$. A learning step with a reactivation was performed directly after the stimuli. The activity of all three hidden layers for both stimuli were measured and give qualitatively similar results.

Supplementary Text

1 Analysis of learning in a single PN

Our neuron model describes the interplay between apical and basal inputs and the resulting effect on synaptic calcium. Still, we have not yet explored how the resulting synaptic plasticity alters sensory processing. We model the modulation of sensory inputs by apical inputs with a simple apical feedback learning circuit Aceituno et al. (2023). This circuit simply compares the firing rate of the neuron with its target firing rate for a specific input, and then sends the difference as a feedback input to the apical dendrite until the neuron has reached its desired and stimulus-specific target activity. Our proof-of-concept toy experiment involves two presynaptic neurons A and B, with synaptic weights w_A and w_B respectively, both connected to a single postsynaptic neuron. Before training, neuron A has a weak connection to the postsynaptic neuron, while neuron B is strongly connected. During training, stimuli S_A and S_B activate neurons A and B, respectively. The feedback circuit now shapes the apical input (i.e., the supervised learning signal) to invert this behavior, effectively swapping the synaptic strengths w_A and w_B. (**Fig. S2A**) shows the evolution of the two synaptic weights over learning. As the neuron increasingly aligns with its target activity, two key metrics indicate the learning progress. First, the amount of top-down apical input required to align the stimulus' response of the neuron with the target decreases over time towards a baseline. Second, the Mean Squared Error (MSE), which corresponds to the squared difference between the output rate of the network and the target value, also reduces (**Fig. S2B**). Finally, we compare for every sample how the gradients (change of activity in the neurons due to the plasticity) compares between the two models (**Fig. S2C**).

Our simple neuron model can thus learn simple tasks both on its rate-based and spike-based version. For the rate model, we can also provide a mathematical argument to show that the change in activity as a result of learning is positively correlated with the change induced by the apical input during learning.

Consider a single postsynaptic neuron with a presynaptic input firing rates denoted by the vector $\mathbf{r}_{\text{pre}}$ that go through synapses with weights $\mathbf{w}$. At a given stimulus presentation, the response of the neuron is given by

$$r^b_{\text{post}} = \phi\left(\langle \mathbf{r}_{\text{pre}}, \mathbf{w}\rangle\right), \tag{33}$$

where $\phi(\cdot)$ is the nonlinear response function which here we take as a rectified linear unit (note that any other monotonically increasing semipositive response function would work).

During learning, the apical input would modify the firing rate,

$$r_{\text{post}} = a\phi\left(\langle \mathbf{r}_{\text{pre}}, \mathbf{w}\rangle\right) \tag{34}$$

and the difference induced during learning by the apical input would be

$$\Delta r_{\text{post}} = r_{\text{post}} - r^b_{\text{post}} = r^b_{\text{post}}(a - a^*). \tag{35}$$

Through the plasticity captured by Eq. 14, after learning we get a new weight vector

$$\mathbf{w}' = \mathbf{w} + \Delta\mathbf{w} = \mathbf{w} + \eta \mathbf{r}_{\text{pre}} r^b_{\text{post}}(a - a^*) \tag{36}$$

which will result in a new response to the stimulus

$$r'_{\text{post}} = \phi\left(\langle \mathbf{r}_{\text{pre}}, \mathbf{w}'\rangle\right) = \phi\left(\langle \mathbf{r}_{\text{pre}}, \mathbf{w}\rangle + \langle \mathbf{r}_{\text{pre}}, \Delta\mathbf{w}\rangle\right). \tag{37}$$

And now by a Taylor expansion we can compute the difference in the responses to stimuli,

$$\Delta r_{\text{post}} = r'_{\text{post}} - r^b_{\text{post}} = \phi'(z)\langle \mathbf{r}_{\text{pre}}, \Delta\mathbf{w}\rangle \tag{38}$$

where $\phi'(z)$ is the derivative of the nonlinear response function evaluated at some presynaptic input in between the inputs before and after the learning. We can plug Eq. 36 in and obtain

$$\Delta r_{\text{post}} = \phi'(z)\langle \mathbf{r}_{\text{pre}}, \eta \mathbf{r}_{\text{pre}} r^b_{\text{post}}(a - a^*)\rangle = \eta\phi'(z) r^b_{\text{post}}(a - a^*)\|\mathbf{r}_{\text{pre}}\|^2. \tag{39}$$

And here we note that the term $r_{\text{post}}(a - a^*)$ already appears in Eq. 35, hence

$$\Delta r_{\text{post}} = \eta\phi'(z)\|\mathbf{r}_{\text{pre}}\|^2 \Delta r_{\text{post}}. \tag{40}$$

Since $\eta,\ \phi'(z),\ \|\mathbf{r}_{\text{pre}}\|^2 \geq 0$, the two changes in postsynaptic firing rate (or neuronal activity) are positively correlated. The only exceptions would be if $\phi'(z) = 0$ or $\mathbf{r}_{\text{pre}} = 0$, which correspond to cases where the neuron is not firing anyway and thus is not involved in representing the stimulus.

2 Formalism for TL and BP

In this section we describe the underlying mathematics of both TL and BP in general terms. Note that practical implementations of either make simplifications or approximations that are useful in practice, which we will not develop in detail. We will use the notation from the main text r^a, r^b to refer to the activities during learning and inference, as those correspond to the activities driven by apical input and at baseline.

A learning algorithm for a neuronal network is an optimization procedure that decreases a loss function

$$\mathcal{L}(\mathbf{w}) = \sum_s l(y_s, f(x_s, \mathbf{w})) \tag{41}$$

were $\mathbf{w}$ are the weights of the connections between neurons (the computational equivalent of synaptic strengths), y_s is a desired output or activity of the network, and $f(\cdot,\cdot)$ is the function implemented by the network parameterized by the weights and the sensory input, and $l(\cdot,\cdot)$ is an error function.

Note that both the error function l and the loss $\mathcal{L}$ are mathematical constructs, which can always be defined for any learning process, but do not need to be implemented by any observable biological quantity. The formalism is still useful to make our analysis understandable and to connect our results to the machine learning literature as well as computational neuroscience frameworks such as predictive coding or energy models, thus we will keep it.

2.1 BP

In classical BP, the change of weights is computed by gradient descent, where the derivatives are computed over the neuron activity. This last technicality is crucial, because it makes BP computationally efficient and it cast it in terms of neuronal activity. Explicitly, this is framed as

$$\Delta \mathbf{w}_{BP} = -\eta \frac{\partial \mathcal{L}(\mathbf{w})}{\partial \mathbf{w}} \tag{42}$$

η is the learning rate. The change for a given weight for one given sample is then expressed as

$$\Delta_{BP} w = -\eta \frac{\partial r_{\text{post}}}{\partial w} \frac{\partial l(y_s, f(x_s, \mathbf{w}))}{\partial r_{\text{post}}} \propto -r^b_{\text{pre}} \frac{\partial l(y_s, f(x_s, \mathbf{w}))}{\partial r_{\text{post}}} \tag{43}$$

where it is common to simplify the notation as $\delta_{BP} = -\dfrac{\partial l(y_s, f(x_s, \mathbf{w}))}{\partial r^b_{\text{post}}}$, giving us

$$\Delta_{BP} w \propto r^{\text{baseline}}_{\text{pre}} \delta_{BP} \tag{44}$$

2.2 TL

In TL, the approach is different in that the loss is minimized by imposing a different activity through an input to each neuron, so that

$$l(y_s, f(x_s, \mathbf{w}, \mathbf{c}^*)) = 0 \tag{45}$$

where $\mathbf{c}^*$ is a vector encoding the signal that imposes the activity onto each neuron. Depending on the architecture, this signal is called a control signal, a top-down prediction or a feedback input, and is computed in various ways (Rao and Ballard, 1999; Meulemans et al., 2022a,b; Song et al., 2024; Scellier and Bengio, 2017). In general, the control is chosen so as to provide the smallest input that will fulfill the aforementioned condition. In energy models, for example, it follows the direction of steepest descent of the loss (Rao and Ballard, 1999; Scellier and Bengio, 2017), although other solutions such as control theory approaches are also used (Meulemans et al., 2021). In the case of unsupervised learning, y_s is simply the activity of the network that is closest to the one imposed by x_s but fulfills a constraint on the activity such as sparsity.

The main point of TL is to provide a plasticity such that the control input will not be required, and thus the network will have learned the right computation without needing this control.

Now a subtle step is to connect the synaptic weight update to the reduction of the control signal $\mathbf{c}$. Since Eq. 45 as an invariant, the change of loss is zero and so is the input, hence

$$l(y_s, f(x_s, \mathbf{w}{+}\Delta\mathbf{w}, \mathbf{c}^*{+}\Delta\mathbf{c})) = 0 \Rightarrow \frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})} \left[\langle \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{c}}, \Delta\mathbf{c} \rangle + \langle \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{w}}, \Delta\mathbf{w} \rangle \right] = 0 \tag{46}$$

where $\langle \cdot, \cdot \rangle$ is a dot product, and the derivatives are taken at for a control signal $\mathbf{c} \in [\mathbf{c}^* + \Delta\mathbf{c}]$ that comes from the mean value theorem. Assuming that the change in control from one iteration to the next will be small (akin to the assumption in BP that the weight updates impose a small change), $\mathbf{c} \approx \mathbf{c}^*$. From it, we can obtain the constraint

$$\langle \frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})} \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{c}}, \Delta\mathbf{c} \rangle = -\langle \frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})} \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{w}}, \Delta\mathbf{w} \rangle \tag{47}$$

where we can already see how the weight update relates to the control signal, which we will now make more precise and in terms that are similar to the standard machine learning formalism.

The first step is to recognize that both terms in Eq. 47 are single scalars, but are computed as dot products of high dimensional quantities. Thus, the constraint is on one degree of freedom, but we want to turn it into an optimization process that applies to many neurons and weights. For this we select the weight update that reduces the control signal in the most efficient way with the minimum weight change.

Starting with the fastest reduction of the control signal. If we want to pick $\Delta\mathbf{c}$ that reduces $\mathbf{c}$ the fastest, we simply set

$$\Delta\mathbf{c} = -\eta\mathbf{c} \tag{48}$$

where the proportionality constant corresponds to the learning rate. Thus, we can rephrase Eq. 47 for single neurons,

$$\eta \frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})} \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial c_{\text{post}}} c_{\text{post}} = \langle \frac{\partial l(y_s, f(x_s, \mathbf{w}, \mathbf{c}))}{\partial f(x_s, \mathbf{w}, \mathbf{c})} \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{w}_{\text{pre}}}, \Delta\mathbf{w}_{\text{pre}} \rangle, \tag{49}$$

and apply a the chain rule to the partial derivatives and make them pass through the postsynaptic activity of the neuron,

$$\frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial c_{\text{post}}} = \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial r_{\text{post}}} \frac{\partial r_{\text{post}}}{\partial c_{\text{post}}}, \quad \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial \mathbf{w}_{\text{pre}}} = \frac{\partial f(x_s, \mathbf{w}, \mathbf{c})}{\partial r_{\text{post}}} \frac{\partial r_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}}. \tag{50}$$

Since there is a derivative that relates the postsynaptic activity to the loss (a scalar to a scalar) on both sides of Eq. 49, we can simplify it to

$$\eta \frac{\partial r_{\text{post}}}{\partial c_{\text{post}}} c_{\text{post}} = \langle \frac{\partial r_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}}, \Delta \mathbf{w}_{\text{pre}} \rangle. \tag{51}$$

Most TL implementations use a control signal is additive (or computed as additive) to the hidden state of the neuron, the feed forward input in standard ANNs. Technically, other choices would work, such as using an additive control to the output, or a multiplicative effect, but to keep our derivations simple we will assume that the additive to the internal state is maintained, and note that the changes would be minimal.

In this line of thought, we introduce an internal neuron variable that conveys the weighted sum of the neuron inputs

$$v_{\text{post}} = \langle \mathbf{r}_{\text{pre}}, \mathbf{w}_{\text{pre}} \rangle \tag{52}$$

which gives the firing rate of the neuron through the relationship $r_{\text{post}} = \phi(v_{\text{post}})$, which we can then use to simplify Eq. 51 to

$$\frac{\partial \phi(v_{\text{post}})}{\partial v_{\text{post}}} \frac{\partial v_{\text{post}}}{\partial c_{\text{post}}} c_{\text{post}} = \frac{\partial \phi(v_{\text{post}})}{\partial v_{\text{post}}} \langle \frac{\partial v_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}}, \Delta \mathbf{w}_{\text{pre}} \rangle \Rightarrow \ \eta c_{\text{post}} = \langle \frac{\partial v_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}}, \Delta \mathbf{w}_{\text{pre}} \rangle. \tag{53}$$

where

$$\frac{\partial v_{\text{post}}}{\partial \mathbf{w}_{\text{pre}}} = \mathbf{r}_{\text{pre}}^{\text{apical}} \Rightarrow \eta c_{\text{post}} = \langle \mathbf{r}_{\text{pre}}^{\text{apical}}, \Delta \mathbf{w}_{\text{pre}} \rangle. \tag{54}$$

where the use of $\mathbf{r}_{\text{pre}}^{\text{apical}}$ as opposed to $\mathbf{r}_{\text{pre}}^{\text{baseline}}$ comes from the fact that the full network is controlled.

Just as with the control signal, there are many possible changes of the presynaptic weight that would reduce the amount of control signal for the simple reason that there are usually many presynaptic neurons. Thus, we assume that the learning weight should be efficient, in the sense that it should provide the maximum decrease of c_{post} for the minimum amount of weight changes $\Delta \mathbf{w}_{\text{pre}}$. Geometrically, the natural way to maximize the effects of a presynaptic weight update is to align the update with the presynaptic inputs, thus setting the angle of the vectors in the dot product to zero.

Taken together, the update rule for a single synapse ends up being

$$\Delta w = \eta r_{\text{pre}}^{\text{apical}} c_{\text{post}} \tag{55}$$

which we can easily recast in the language of machine learning as

$$\Delta w \propto r_{\text{pre}}^{\text{apical}} \delta_{TL}. \tag{56}$$

Note that a crucial component of TL is that r_{pre}^{a} is taken when the network is controlled, thus it is the presynaptic activity as it will be at target (when the network has already learned).

2.3 Difference in synaptic updates

Contrasting the learning rules from Eq. 44 and Eq. 56, we find that both contain the presynaptic activity and a δ term. The later one can be used interchangeably between both approaches (see **Supp. 3**), but the presynaptic term differs in that r_{pre}^{a} for TL is the same as the activity after learning, while in BP, r_{pre}^{b} is the network activity during inference.

3 Links between BP and TL

The two families of algorithms that we study are not separated classes, but rather two extremes of a spectrum. While both algorithms adapt the synaptic weights to modify the activity so that a desired output is provided, BP uses the neuronal responses of the network to a stimulus, while TL disregards the current responses and uses instead the desired neuronal responses. In this section we express this statement in mathematical terms.

The intuitive way to understand where the two algorithms intersect is to consider a small control input $\mathbf{c}$. In this case, the activity of the network with or without the control is very similar, since the modifications due to the control are, by definition, small. In fact, there are works which explicitly use architectures that were originally designed for TL implementations (Rao and Ballard, 1999), but were then turned into BP implementations by using a weak control that is similar to BP(Scellier and Bengio, 2017; Whittington and Bogacz, 2017).

This naturally implies that the difference between both algorithms fades away in the late stages of learning, where both desired and current neuronal responses are very close and thus both algorithms can be very similar. Since our main goal is to differentiate them, our analysis concerns the learning process at the beginning of learning, where the network has just been initialized and the control signal in TL would have to be strong.

In the next sections we prove that TL and BP are similar by showing analytically that BP with strong teaching signals that alter the neuron activities results in a learning procedure that is better described by TL than to BP, while TL with weak control is better described by BP than by TL.

3.1 BP with strong apical inputs is a close approximation of TL

At the single neuron level, the learning rule from BP is

$$\Delta_{BP} w = -\eta r_{\text{pre}}^{\text{baseline}} \delta_{BP} \tag{57}$$

where $r_{\text{pre}}^{\text{baseline}}$ is the presynaptic activity during the feedforward pass, implying that either the apical input would be very close to baseline to avoid altering the firing rate on the presynaptic neurons. However, if we allow the apical input to have a strong driving effect as our model and previous literature suggest (Larkum et al., 1999), we would have a different learning procedure. Such a learning procedure is actually closer to TL than to BP in terms of synaptic updates and neuronal activity during learning (as measured by cortical reactivations).

The approach that we follow is to show that the matrix transpose at the early learning stages is very similar to the matrix inverse which emerges from δ_{TL}, and thus using it to compute a strong feedback input is very similar to using feedback to control the network and use the control signal for learning.

First we outline some general properties that we would expect from the matrix $\frac{\partial \mathbf{r}_L}{\partial \mathbf{r}}$. At the beginning of learning, which is where the alignment between reactivations and stimulus differs for BP and TP, the network is not trained to that specific stimulus-output pair, and thus every entry of the matrix can be considered as random. Intuitively, if we sampled various networks, we would expect different entries but with similar statistics; thus it is natural to treat any of those matrix entries as drawn independently from a given probability distribution. The second property is that we expect the effects of all neurons on the output to be balanced, meaning that the synaptic weights (and the derivatives that emerge from them) have an average of zero. We will also assume that the network is large, implying that the matrix has many elements, and that the variance of the distribution is not infinite.

The inverse of a matrix M is defined simply by the property that

$$MM^{-1} = I \tag{58}$$

where I is the identity matrix. Computing an exact inverse is often a complex procedure with a complexity of $O(n^3)$, where n is the number of rows. However, for large matrices whose entries are independently drawn from a probability distribution that has zero expectation $\mu_M = 0$ (in the language of computational neuroscience, this corresponds to a balanced distribution) and with finite variance $\sigma_M < \infty$, we can obtain a good approximation by using the transpose $M^\top$. To illustrate this, consider the value of entries in the diagonal and off diagonal of the product $P = MM^\top$

$$\mathbf{E}_i \left[P_{ii}\right] = \mathbf{E}_i \left[\sum_k M_{ik} M_{ik}\right] = n\mathbf{E}_{i,k} \left[M_{ik}^2\right] \tag{59}$$

$$\mathbf{E}_{i\neq j} \left[P_{ij}\right] = \mathbf{E}_{i\neq j} \left[\sum_k M_{ik} M_{jk}\right] = n\mathbf{E}_{i,j,k,l} \left[M_{ij} M_{kl}\right] \tag{60}$$

and by using the standard rules to compute the variance and expectation of the product of independent variables,

$$\mathbf{E}_i \left[P_{ii}\right] = n\sigma_M^2 \tag{61}$$

$$\mathbf{E}_{i\neq j} \left[P_{ij}\right] = n\mu_M^2 = 0 \tag{62}$$

$$\mathbf{Var}_i \left[P_{ii}\right] = n\sigma_M^2 \tag{63}$$

$$\mathbf{Var}_{i\neq j} \left[P_{ij}\right] = n\sigma_M^2 \tag{64}$$

Thus, the matrix P is a matrix with a diagonal $n\sigma_M^2 \pm \sqrt{n}\sigma_M$ and off-diagonal entries with mean zero and standard deviation of $\sqrt{n}\sigma_M$.

Importantly, the matrix is only close to TL if $MM^{-1} = I$, implying that we need to scale the matrix M such that $\sigma_M = n^{-1/2}$, when n is large we get that $P \approx I$, hence $M^-1 \approx M^\top$. In this calculation, we have to pick σ_M, but in the network case we might not have this freedom as the variance at initialization is crucial for learning. Instead, we can adapt the scaling of the control input δ such that it cancels out σ_M.

3.2 Computationally efficient TL with weak feedback is an approximation of BP

Having shown that BP with strong feedback is close to TL, we now prove the complementary direction and argue that TL with weak feedback is actually closer to BP than to TL. For simplicity and to align with our implementation of TL, we will use the nomenclature of control theory, although a similar argument has been build for other specific implementations such as predictive coding (Whittington and Bogacz, 2017) or Equilibrium Propagation (Scellier and Bengio, 2017).

As per our previous argument, we note that BP and TL are fundamentally different in that the presynaptic activity is altered during TL synaptic updates,

$$\Delta_{TL} w \propto -r_{\text{pre}}^{\text{apical}} \delta_{TL} \tag{65}$$

where $r_{\text{pre}}^{\text{apical}}$ is the presynaptic activity when the control signal is provided to the network. An important point here is that the control signal is used as δ_{TL}.

Naturally, if we have a very weak control with respect to the neuronal activity, then $r_{pre}^a \approx r_{pre}^b$, hence the main problem is what is the control signal. Usually, controlling a system requires pushing it until it reaches a certain goal. To have a weak control signal would then mean to either not reach the target (which would then not be TL as per our definition) or set a target that is not the correct output, but rather a nudge in the direction of the target. The latter choice leads to

$$\mathbf{r}_L \xrightarrow[\text{wTL}]{} \mathbf{r}_L - \eta \mathbf{e} \Rightarrow \mathbf{r} \to \mathbf{r} + \frac{\partial \mathbf{r}_L}{\partial \mathbf{r}} \delta_{TL} \tag{66}$$

where $\eta \ll 1$ is the strength of the nudge, and the use of a derivative on the last term is valid because the changes in activity are very small. Given that the number of neurons is usually much higher than the dimensionality of the output, there are many possible configurations of δ_{TL}. The one that minimizes the norm of δ_{TL} on average across all errors would be the Moore-Penrose pseudoinverse

$$\delta_{TL}^{MP} = \eta \left(\frac{\partial \mathbf{r}_L}{\partial \mathbf{r}} \right)_{MP}^{-1} \mathbf{e} \tag{67}$$

which is, like other matrix inverses, hard to compute (although valid). A simpler approach would be to simply evaluate the effect that each neuron has on the output and simply increase the control to each neuron until the target is achieved. This would imply that for neuron n we would provide an input proportional to the derivative of the output with respect to that neuron,

$$\delta_{TL}^{*} = \eta \left(\frac{\partial \mathbf{r}_L}{\partial \mathbf{r}} \right)^{\top} \mathbf{e}. \tag{68}$$

Which aligns with BP.

4 Somatic calcium imaging

While we could not image intracellular calcium dynamics during plasticity induction events, we reproduced typical induction events in a subset of neurons (n = 7 neurons, 4 mice) after post-induction EPSP recordings were complete. During these stimulations, we recorded voltage and calcium signals for activity patterns similar to the ones we used for plasticity induction. We found a significant correlation between somatic calcium recordings (derived as ΔF/F AUC) and PSST-AUC (**Fig. S6B**, $R_{Spearman} = 0.8303,\ p = 0.0029$).

5 Apical plasticity

In contrast to our results on basal synapses, we found that for apical synapses neither the plasticity induction AUC ($R_{Spearman} = 0.1049,\ p = 0.7456$, **Fig. S7D**) nor the number of APs ($R_{Spearman} = 0.3996,\ p = 0.1982$, **Fig. S7E**) are good predictors for changes in synaptic strength. This difference in the predictability of basal versus apical plasticity is also reflected in the finding, that overall changes in basal EPSP amplitudes did not correlate with changes in apical EPSPs ($R_{Spearman} = 0.2308,\ p = 0.4705$, **Fig. S7F**).

Parameter	Value	Reference(s)
Duration of AP	1 to 2ms	(Gulledge and Stuart, 2003)
Refractory period	6ms	(Gulledge and Stuart, 2003)
Resting membrane potential	-78 to -74 mV	Own experiments
Peak AP	55 to 65mV	(Kasper et al., 1994)
Number of spikes in burst	2 to 4 spikes	Own experiments
Duration of EPSP	20 to 30ms	Own experiments
bAP half-width	1 to 4 ms	(Stuart et al., 1997)
LTD/LTP threshold	$\theta_{LTP} > \theta_{LTD}$ $\theta_{LTP} \approx 2 \cdot \theta_{LTD}$	(Graupner and Brunel, 2012), (Nevian and Sakmann, 2006)
Calcium spike threshold	2.28 (± 0.14) nA	(Larkum et al., 1999)
Initiation of plateau potential	5 to 10 ms	(Larkum et al., 1999)
Duration of plateau potential	50 to 100 ms	Own experiments
Amplitude of plateau potential	15 mV	(Larkum et al., 1999), Own experiments
NMDAR influx vs voltage	$1/(1 + \exp(-0.22 * (x + 49)))$	(McRory et al., 2001), (Slutsky et al., 2004)
Calcium exp. time constant	81 ± 13 ms	(Majewska et al., 2000)
Maximal calcium in synapse	3 a.u. (see NMDAR influx)	n/a
Basal-apical coupling window	20 to 30 ms	(Larkum et al., 2004)

Table 1: Hyperparameters of the neuron model with their associated value and references.

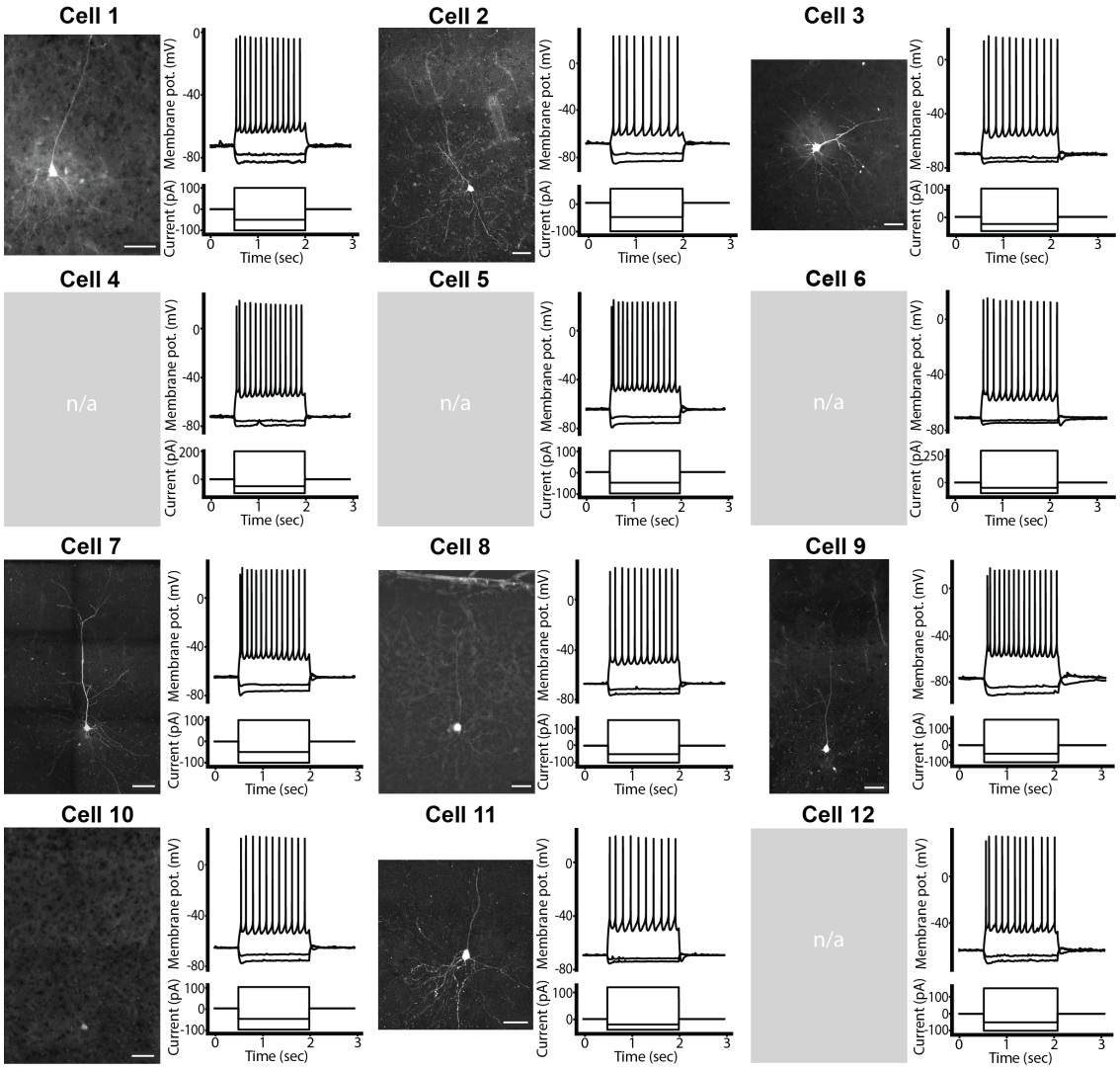

Figure S1: Biocytin-Streptavidin stained neurons (left) and typical hyperpolarizations and action potential trains (right top) induced by negative and positive current injection steps (right bottom) from recorded neurons. For cells 3, 4, 5, and 11, no Biocytin-Streptavidin staining/morphology reconstruction was performed.

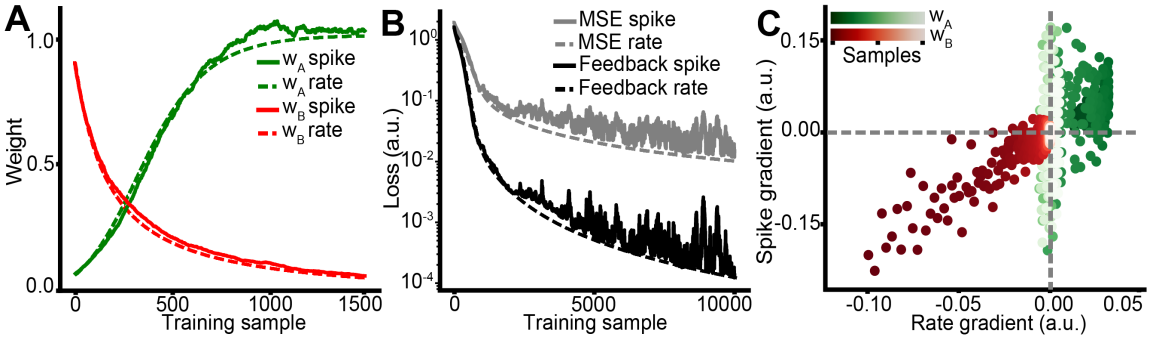

Figure S2: Comparison of spike-based versus rate-based neuron model learning dynamics. (**A**) Training of the synaptic weights. (**B**) Evolution of the magnitude of the feedback signal and loss function. (**C**) Comparison of the change in gradients across the simulation.

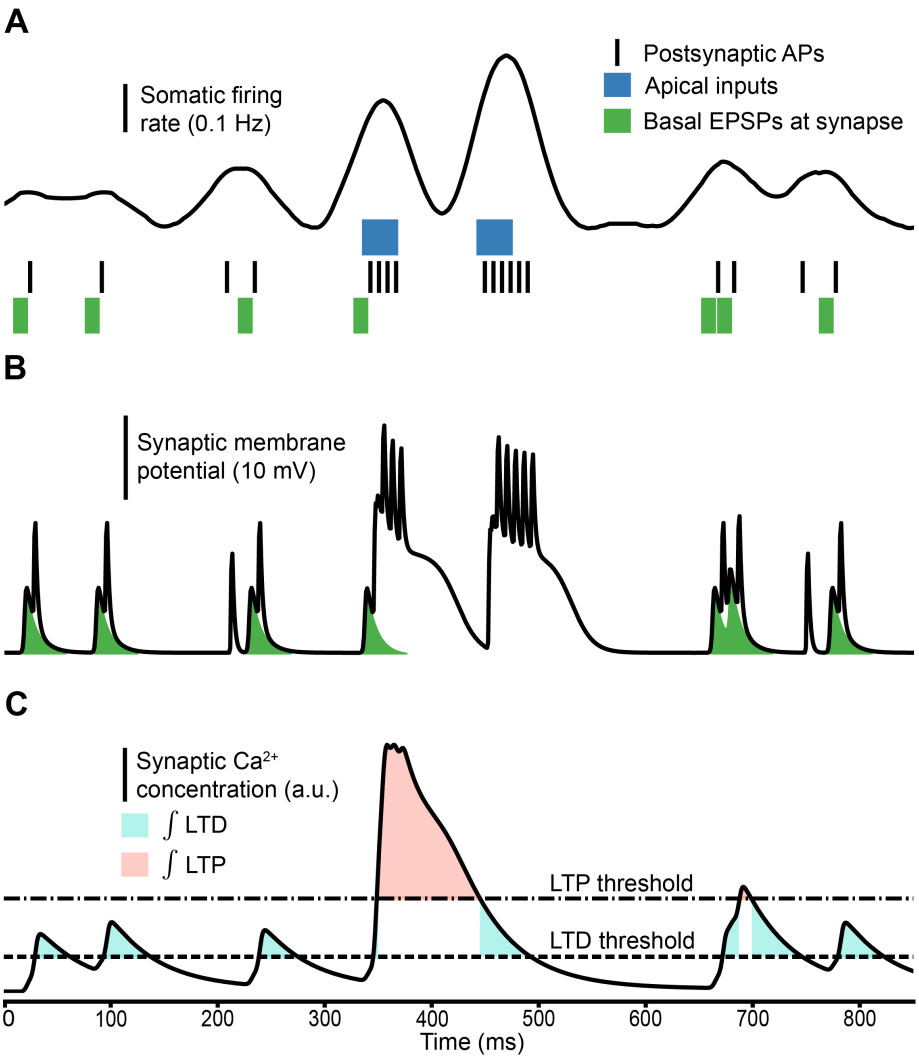

Figure S3: Effects of apical input on somatic firing rate, synaptic membrane potential, calcium influx, and plasticity. (**A**) Neuron firing rate modulated by apical input. (**B**) Synaptic membrane potential integrating EPSPs, bAPs, and plateau potentials. (**C**) Synaptic calcium concentrations with LTD and LTP thresholds and their integration.

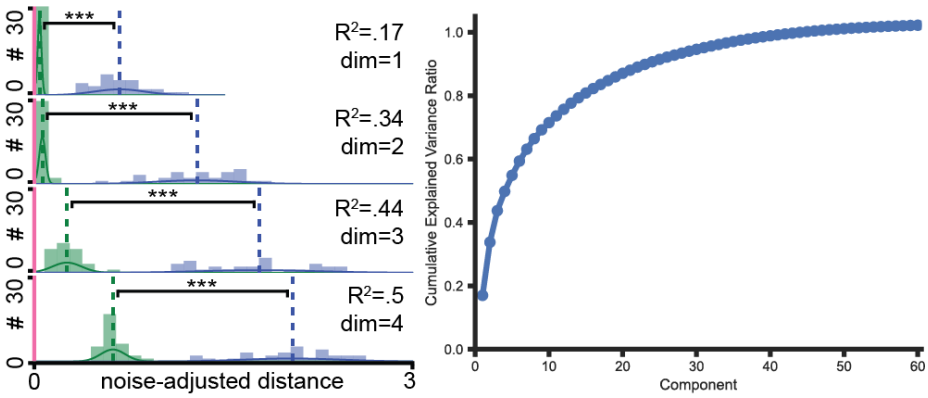

Figure S4: Significance of BP and TL hypotheses across principle components. BP and TL hypotheses for the first (early-to-late vector) to the first four principle components, with explained variance for every additional component (left). The cumulative explained variance ratio for all principle components (right).

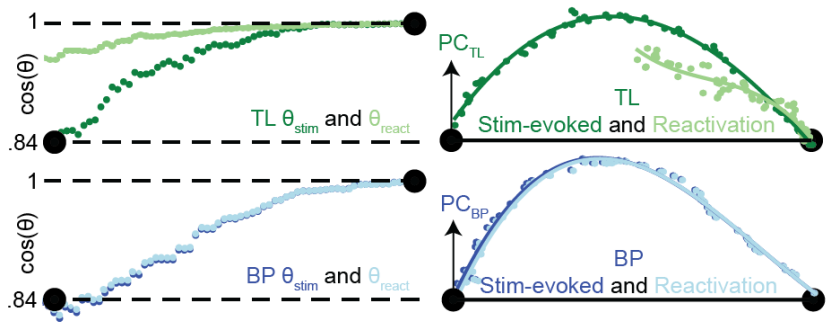

Figure S5: Angles and Orthogonal PC of BP and TL training within ANNs. Cosine similarity of the stimulus-evoked responses and reactivations for the BP and TL network simulations to the late-training centroid (left). The movement of the trajectories in the subspace comprising of the early-to-late training vector and the first principle component orthogonal to this vector (right).

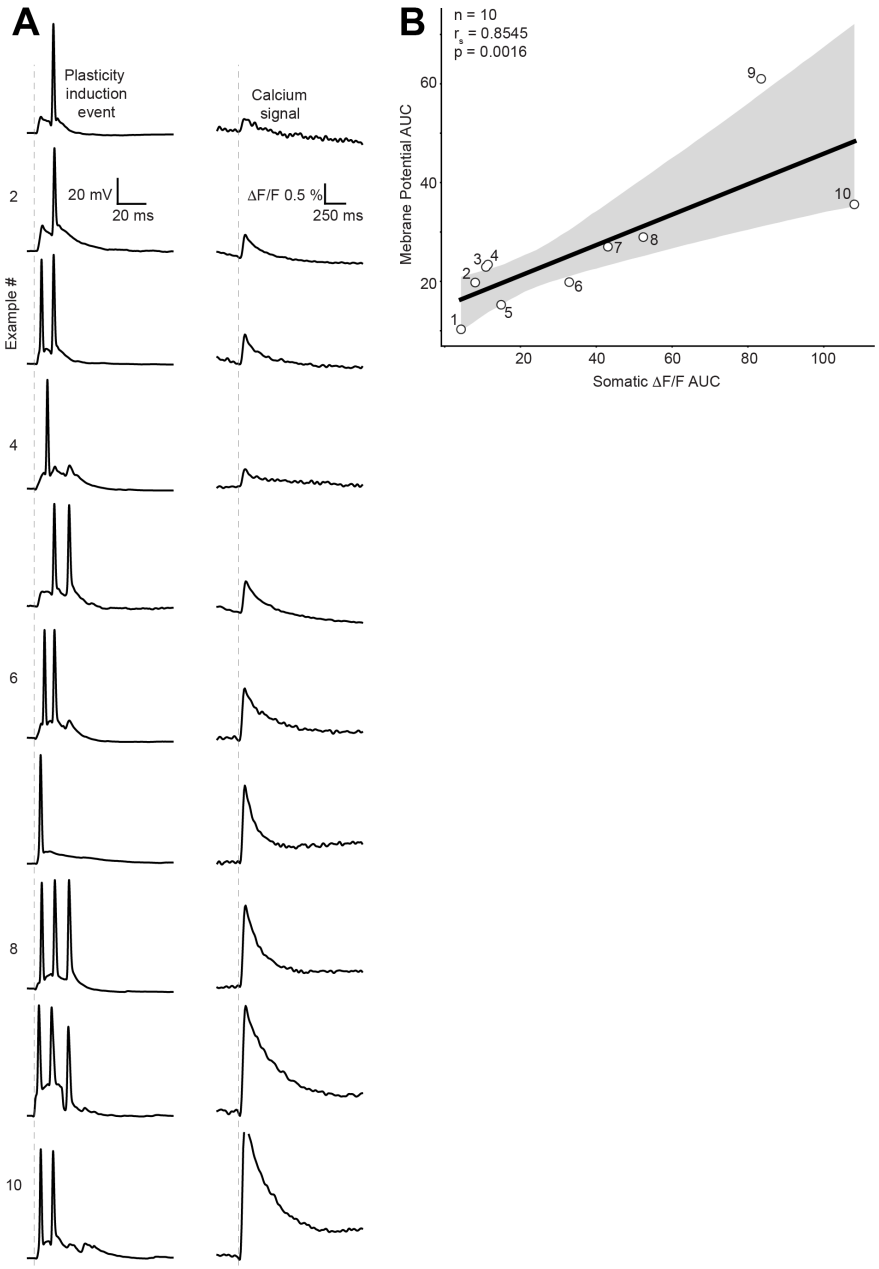

Figure S6: (**A**) Somatic membrane potential recording during plasticity induction (left) and somatic calcium ΔF/F during the same event (right). In total, showing events from 10 recordings in 6 neurons (4 mice) from **Fig. S1**. (**B**) Linear fit of induction event AUCs and somatic calcium signal AUCs showing significant correlation (n=6, 10 events total, R_S=0.8303, p=0.0029)

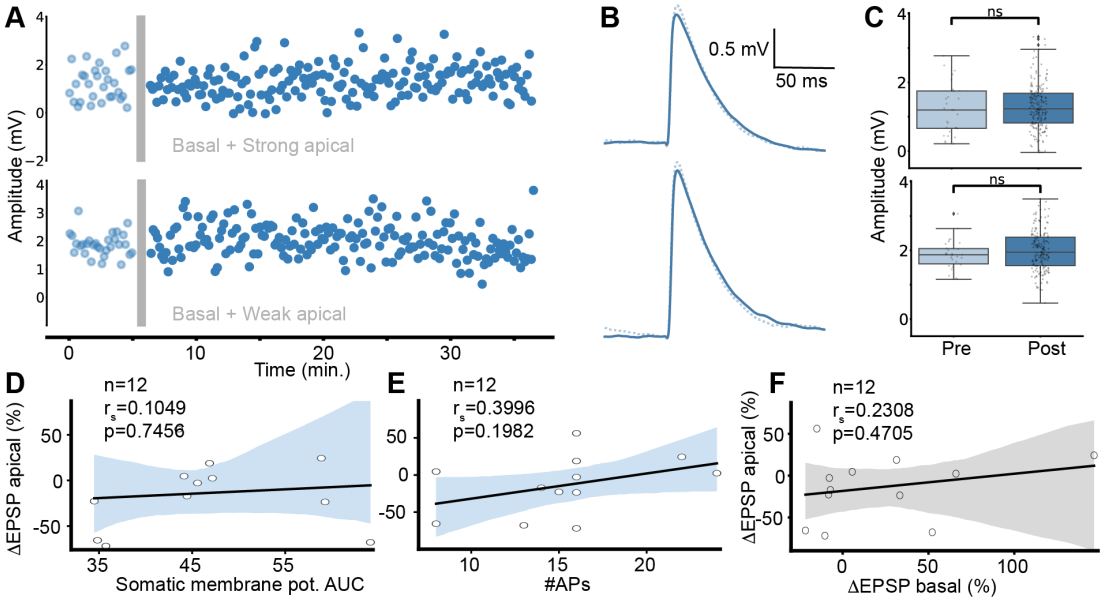

Figure S7: Representative example recordings of EPSP amplitudes during apical stimulation before and after the plasticity induction event (**A**) EPSP amplitude evolution over time; each dot represents one EPSP pre (light blue) and post (dark blue) plasticity induction event (gray bar). (**B**) The quantification of EPSPs for two sample cells shows no significant change in EPSP amplitudes. (**C, D**) Linear correlation analysis between changes in apical EPSP amplitude and somatic membrane potential AUC ((C) n=12, r_S=0.1049, p=0.7456) and number of APs ((D) n=12, r_S=0.3996, p=0.1982), respectively. (**E**) Correlation analysis of basal and apical ΔEPSP (n=12, r_S=0.3669, p=0.2408). (**F**) Correlation between basal and apical EPSP amplitude changes $r_s = 0.2308,\ p = 0.4705$.