Towards deep learning with segregated dendrites

Jordan Guergiuev[1,2], Timothy P. Lillicrap[4], and Blake A. Richards[1,2,3,*]

1 Department of Biological Sciences, University of Toronto Scarborough, Toronto, ON, Canada
2 Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada
3 Learning in Machines and Brains Program, Canadian Institute for Advanced Research, Toronto, ON, Canada
4 DeepMind Technologies Inc., London, UK

*** Corresponding author, email: blake.richards@utoronto.ca**

Abstract

Deep learning has led to significant advances in artificial intelligence, in part, by adopting strategies motivated by neurophysiology. However, it is unclear whether deep learning could occur in the real brain. Here, we show that a deep learning algorithm that utilizes multi-compartment neurons might help us to understand how the brain optimizes cost functions. Like neocortical pyramidal neurons, neurons in our model receive sensory information and higher-order feedback in electrotonically segregated compartments. Thanks to this segregation, the neurons in different layers of the network can coordinate synaptic weight updates. As a result, the network can learn to categorize images better than a single layer network. Furthermore, we show that our algorithm takes advantage of multilayer architectures to identify useful representations—the hallmark of deep learning. This work demonstrates that deep learning can be achieved using segregated dendritic compartments, which may help to explain the dendritic morphology of neocortical pyramidal neurons.

Introduction

Deep learning refers to an approach in artificial intelligence (AI) that utilizes neural networks with multiple layers of processing units. Importantly, deep learning algorithms are designed to take advantage of these multi-layer network architectures in order to generate hierarchical representations wherein each successive layer identifies increasingly abstract, relevant variables for a given task (Bengio and LeCun, 2007; LeCun et al., 2015). In recent years, deep learning has revolutionized machine learning, opening the door to AI applications that can rival human capabilities in pattern recognition and control (Mnih et al., 2015; Silver et al., 2016; He et al., 2015). Interestingly, the representations that deep learning generates resemble those observed in the neocortex, particularly in higher-order sensory areas (Kubilius et al., 2016; Khaligh-Razavi and Kriegeskorte, 2014; Cadieu et al., 2014), suggesting that something akin to deep learning is occurring in the mammalian brain (Yamins and DiCarlo, 2016; Marblestone et al., 2016).

Yet, a large gap exists between deep learning in AI and our current understanding of learning and memory in neuroscience. In particular, unlike deep learning researchers, neuroscientists do not yet have a solution to the "credit assignment problem" (Rumelhart et al., 1986; Lillicrap et al., 2016; Bengio et al., 2015). The credit assignment problem refers to the fact that the behavioral consequences of synaptic changes in early layers of a neural network depend on the current connections in the downstream layers. For example, consider the behavioral effects of synaptic changes, i.e. long-term potentiation/depression

(LTP/LTD), occurring between different sensory circuits of the brain. Exactly how these synaptic changes will impact behavior and cognition depends on the downstream connections between the sensory circuits and motor or associative circuits (Figure 1A). Hence, learning to optimize some behavioral or cognitive function requires a method for identifying what a given neuron's place is within the wider neural network, i.e. it requires a means of assigning "credit" (or "blame") to neurons in early sensory areas for their contribution to the final behavioral output (LeCun et al., 2015; Bengio et al., 2015). Despite its importance for real-world learning, the credit assignment problem has received little attention in neuroscience.

The lack of attention to credit assignment in neuroscience is, arguably, a function of the history of biological studies of synaptic plasticity. Due to the well-established dependence of LTP and LTD on presynaptic and postsynaptic activity, current theories of learning in neuroscience tend to emphasize Hebbian learning algorithms (Dan and Poo, 2004; Martin et al., 2000), that is, learning algorithms where synaptic changes depend solely on presynaptic and postsynaptic activity. Hebbian learning models can produce representations that resemble the representations in the real brain (Zylberberg et al., 2011; Leibo et al., 2017) and they are backed up by decades of experimental findings (Malenka and Bear, 2004; Dan and Poo, 2004; Martin et al., 2000). But, current Hebbian learning algorithms do not solve the credit assignment problem, nor do global neuromodulatory signals used in reinforcement learning (Lillicrap et al., 2016). As a result, deep learning systems from AI that can do credit assignment outperform existing Hebbian models of perceptual learning on a variety of tasks (Yamins and DiCarlo, 2016; Khaligh-Razavi and Kriegeskorte, 2014). This suggests that a critical, missing component in our current models of the neurobiology of learning and memory is an explanation of how the brain solves the credit assignment problem.

However, the most common solution to the credit assignment problem in AI is to use the back-propagation of error algorithm (Rumelhart et al., 1986). Backpropagation assigns credit by *explicitly* using current downstream synaptic connections to calculate synaptic weight updates in earlier layers, commonly termed "hidden layers" (LeCun et al., 2015) (Figure 1B). This technique, which is sometimes referred to as "weight transport", involves non-local transmission of synaptic weight information between layers of the network (Lillicrap et al., 2016; Grossberg, 1987). Weight transport is clearly unrealistic from a biological perspective (Bengio et al., 2015; Crick, 1989). It would require early sensory processing areas (e.g. V1, V2, V4) to have precise information about *billions* of synaptic connections in downstream circuits (MT, IT, M2, EC, etc.). According to our current understanding, there is no physiological mechanism that could communicate this information in the brain. Some deep learning algorithms utilize purely Hebbian rules (Scellier and Bengio, 2016; Hinton et al., 2006). But they depend on feedback synapses that are symmetric with feedforward synapses to solve the credit assignment problem (Scellier and Bengio, 2016; Hinton et al., 2006), which is essentially a version of weight transport. Altogether, these artificial aspects of current deep learning solutions to credit assignment have rendered many scientists skeptical to the proposal that deep learning occurs in the real brain (Crick, 1989; Grossberg, 1987; Harris, 2008; Urbanczik and Senn, 2009).

Recent findings have shown that these problems may be surmountable, though. Lillicrap et al. (2016), Lee et al. (2015) and Liao et al. (2015) have demonstrated that it is possible to solve the credit assignment problem even while avoiding weight transport or symmetric feedback weights. The key to these learning algorithms is the use of feedback signals from output layers that are transmitted to calculate a local error signal, which then guides synaptic updates in hidden layers (Lee et al., 2015; Lillicrap et al., 2016; Liao et al., 2015). These learning algorithms can take advantage of multi-layer architectures, leading to performance that rivals backpropagation (Lee et al., 2015; Lillicrap et al., 2016; Liao et al., 2015). Hence, this work has provided a significant breakthrough in our understanding of how the real brain might do credit assignment.

Nonetheless, the models of Lillicrap et al. (2016), Lee et al. (2015) and Liao et al. (2015) involve some problematic assumptions. Specifically, although it is not directly stated in all of the papers, there is an implicit assumption that there is a separate feedback pathway for transmitting the signals that drive synaptic updates in hidden layers (Figure 2A). Such a pathway is required in these models because the synaptic weight updates in the hidden layers depend on the difference between feedback that is

generated in response to a purely feedforward propagation of sensory information, and feedback that is guided by a teaching signal (Lillicrap et al., 2016; Lee et al., 2015; Liao et al., 2015). In order to calculate this difference, sensory information must be transmitted *separately* from the feedback signals that are used to drive learning. In single compartment neurons, keeping feedforward sensory information separate from feedback signals is impossible without a separate pathway. At face value, such a pathway is possible. But, closer inspection uncovers a couple of difficulties with such a proposal.

First, the error signals that solve the credit assignment problem are not global error signals (like neuromodulatory signals used in reinforcement learning). Rather, they are *cell-by-cell* error signals. This would mean that the feedback pathway would require some degree of pairing, wherein each neuron in the hidden layer is paired with a feedback neuron (or circuit). That is not impossible, but there is no evidence to date of such an architecture in the neocortex. Second, the error that is communicated to the hidden layer is signed (i.e. it can be positive or negative), and the sign determines whether LTP or LTD occur in the hidden layer neurons (Lee et al., 2015; Lillicrap et al., 2016; Liao et al., 2015). Communicating signed signals with a spiking neuron can theoretically be done by using a baseline firing rate that the neuron can go above (for positive signals) or below (for negative signals). But, in practice, such systems are difficult to operate because as the error gets closer to zero any noise in the spiking of the neuron can switch the sign of the signal, which switches LTP to LTD, or *vice versa*. This means that as learning progresses the network's ability to communicate error signs gets *worse*. Therefore, the real brain's specific solution to the credit assignment problem is unlikely to involve a separate feedback pathway for cell-by-cell, signed signals to instruct plasticity.

However, segregating the integration of feedforward and feedback signals does not require a separate pathway if neurons have more complicated morphologies than the point neurons typically used in artificial neural networks. Taking inspiration from biology, we note that real neurons are much more complex than single-compartments, and different signals can be integrated at distinct dendritic locations. Indeed, in the primary sensory areas of the neocortex, feedback from higher-order areas arrives in the distal apical dendrites of pyramidal neurons (Manita et al., 2015; Budd, 1998; Spratling, 2002), which are electrotonically very distant from the basal dendrites where feedforward sensory information is received (Larkum et al., 1999, 2007, 2009). Thus, as has been noted by previous authors (Körding and König, 2001; Spratling, 2002; Spratling and Johnson, 2006), the anatomy of pyramidal neurons may actually provide the necessary segregation of feedforward and feedback information to calculate local error signals and perform deep learning in biological neural networks.

Here, we show how deep learning can be implemented if neurons in hidden layers contain segregated "basal" and "apical" dendritic compartments for integrating feedforward and feedback signals separately (Figure 2B). Our model builds on previous neural networks research (Lee et al., 2015; Lillicrap et al., 2016) as well as computational studies of supervised learning in multi-compartment neurons (Urbanczik and Senn, 2014; Körding and König, 2001; Spratling and Johnson, 2006). Importantly, we use the distinct basal and apical compartments in our neurons to integrate feedback signals separately from feedforward signals. With this, we build a local target for each hidden layer that coordinates learning across the layers of the network. We demonstrate that even with random synaptic weights for feedback into the apical compartment, our algorithm can coordinate learning to achieve classification of the MNIST database of hand-written digits better than can be achieved with single layer networks. Furthermore, we show that our algorithm allows the network to take advantage of multi-layer structures to build hierarchical, abstract representations, one of the hallmarks of deep learning (LeCun et al., 2015). Our results demonstrate that deep learning can be implemented in a biologically feasible manner if feedforward and feedback signals are received at electrotonically segregated dendrites, as is the case in the mammalian neocortex.

Results

A network architecture with segregated dendritic compartments

Deep supervised learning with local weight updates requires that each neuron receive signals that can be used to determine its "credit" for the final behavioral output. We explored the idea that the

cortico-cortical feedback signals to pyramidal cells could provide the required information for credit assignment. In particular, we were inspired by four observations from both machine learning and biology:

1. Current solutions to credit assignment without weight transport require segregated feedforward and feedback signals (Lee et al., 2015; Lillicrap et al., 2016).

2. In the neocortex, feedforward sensory information and higher-order cortico-cortical feedback are largely received by distinct dendritic compartments, namely the basal dendrites and distal apical dendrites, respectively (Spratling, 2002; Budd, 1998).

3. The distal apical dendrites of pyramidal neurons are electrotonically distant from the soma, such that passive transmission from the distal dendrites to the soma is significantly attenuated. Instead, apical communication to the soma depends on active propagation through the apical dendritic shaft driven by voltage-gated calcium channels. These non-linear, active events in the apical shaft generate prolonged upswings in the membrane potential, known as "plateau potentials", which can drive burst firing at the soma (Larkum et al., 1999, 2009).

4. Plateau potentials driven by apical activity can guide plasticity *in vivo* (Bittner et al., 2015).

With these considerations in mind, we hypothesized that the computations required for credit assignment could be achieved without any separate pathways for feedback signals. Instead, they could be achieved by having two distinct dendritic compartments in each hidden layer neuron: a "basal" compartment, strongly coupled to the soma for integrating feedforward information, and an "apical" compartment for integrating feedback information that would only drive activity at the soma when "plateau potentials" occur (Figure 3A).

As an initial test of this concept we built a network with a single hidden layer. Although this network is not very "deep", even a single hidden layer can improve performance over a one-layer architecture if the learning algorithm solves the credit assignment problem (Bengio and LeCun, 2007; Lillicrap et al., 2016). Hence, we wanted to initially determine whether our network could take advantage of a hidden layer to reduce error at the output layer.

The network architecture is illustrated in Figure 3A. An image from the MNIST data set is used to set the spike rates of $\ell = 784$ Poisson point-process neurons in the input layer (one neuron per image pixel, rates-of-fire determined by pixel intensity). These project to a hidden layer with $m = 500$ neurons. The neurons in the hidden layer are composed of three distinct compartments with their own voltages: the apical compartments (with voltages described by the vector $\boldsymbol{A}(t) = [A_1(t), ..., A_m(t)]$), the basal compartments (with voltages $\boldsymbol{B}(t) = [B_1(t), ..., B_m(t)]$), and the somatic compartments (with voltages $\boldsymbol{C}(t) = [C_1(t), ..., C_m(t)]$). (*Note*: for notational clarity, all vectors and matrices in the paper are in boldface.) The voltages in the dendritic compartments are calculated as weighted sums of a postsynaptic potential kernel convolved with the incoming spike train (see Materials and Methods, equations (10) and (11)), while the somatic voltages are calculated as leaky integrators of the dendritic inputs, i.e. for the i^{th} hidden layer neuron:

$$\frac{dC_i(t)}{dt} = -g_L C_i(t) + g_B(B_i(t) - C_i(t)) + g_A(A_i(t) - C_i(t)) \tag{1}$$

where g_L, g_B and g_A represent the leak conductance, the conductance from the basal dendrites, and the conductance from the apical dendrites, respectively. Note, for mathematical simplicity we are assuming a resting membrane potential of 0 V and a membrane capacitance of 1 F (these values do not affect the results). We implement electrotonic segregation in the model by altering the g_A value—low values for g_A lead to electrotonically segregated apical dendrites. In the initial set of simulations we set $g_A = 0$, but we relax this hard constraint in later simulations. The somatic compartments generate spikes using Poisson processes. The instantaneous rates of these processes are described by the vector

$\boldsymbol{\lambda}^C(t) = [\lambda_1^C(t), ..., \lambda_m^C(t)]$, which is in units of spikes/s or Hz. These rates-of-fire are determined by a non-linear sigmoid function, $\sigma(\cdot)$, applied to the somatic voltages, i.e. for the i^th hidden layer neuron:

$$\begin{aligned} \lambda_i^C(t) &= \lambda_{max}\sigma(C_i(t)) \\ &= \lambda_{max}\frac{1}{1+e^{-C_i(t)}} \end{aligned} \tag{2}$$

where λ_{max} is the maximum rate-of-fire for the neurons.

Spiking inputs from the input layer arrive at the basal compartments via the $m \times \ell$ synaptic weight matrix $\boldsymbol{W}^0$, hidden layer somatic spikes are projected to the output layer neurons via the $n \times m$ synaptic weight matrix $\boldsymbol{W}^1$, and spiking inputs from the output layer arrive at the apical compartments via the $m \times n$ synaptic weight matrix $\boldsymbol{Y}$ (Figure 3A).

The output layer neurons consist of $n = 10$ two compartment neurons (one for each image category), similar to those used in a previous model of dendritic prediction learning (Urbanczik and Senn, 2014). The output dendritic voltages ($\boldsymbol{V}(t) = [V_1(t), ..., V_n(t)]$) and somatic voltages ($\boldsymbol{U}(t) = [U_1(t), ..., U_n(t)]$) are updated in a similar manner to the hidden layer basal compartment and soma, respectively (see Materials and Methods, equations (14) and (15)). As well, like the hidden layer neurons, the output neurons spike using Poisson processes whose instantaneous rates, $\boldsymbol{\lambda}^U(t) = [\lambda_1^U(t), ..., \lambda_n^U(t)]$, are determined by the somatic voltages, i.e. $\lambda_i^U(t) = \lambda_{max}\sigma(U_i(t))$. Unlike the hidden layer, however, the output layer neurons also receive a set of "teaching signals", specifically, excitatory and inhibitory conductances that push their voltages towards target values (see Materials and Methods, equations (15) and (16)). These teaching signals are similar to the targets that are used for training in deep artificial neural networks (Bengio and LeCun, 2007; LeCun et al., 2015). Whether any such teaching signals exist in the real brain is unknown, though there is evidence that animals can represent desired behavioral outputs with internal goal representations (Gadagkar et al., 2016).

Critically, we define two different modes of integration in the hidden layer neurons: "transmit" and "plateau". During the transmit mode, the apical compartment is considered to be electrotonically segregated from the soma so it affects the somatic voltage minimally (depending on g_A), and most of the conductance to the soma comes from the basal compartment (Figure 3B, left). In contrast, during the plateau mode, the apical voltage is averaged over the most recent 20-30 ms period and the sigmoid non-linearity is applied to it, giving us "plateau potentials" (see equation (3) below). These non-linear versions of the apical voltages are then transmitted to the somatic and basal compartments for synaptic updates (Figure 3B, right). The intention behind this design was to mimic the non-linear transmission from the apical dendrites to the soma that occurs during a plateau potential driven by calcium spikes in the apical dendritic shaft (Larkum et al., 1999). Furthermore, the temporal averaging was intended to mimic, in part, the temporal dynamics introduced by NMDA plateaus (Schiller et al., 2000), which are particularly good for driving apical calcium spikes (Larkum et al., 2009).

To train the network we alternate between two phases. First, we present an image to the input layer without any signals to the output layer during the "forward" phase, which occurs between times t_0 to t_1. At t_1 a plateau potential occurs in all the hidden layer neurons and the "target" phase begins. During this phase the image continues to drive the input layer, but now the output layer also receives teaching conductances, which force the output units closer to the correct answer. For example, if an image of a '9' is presented, then over the time period t_1-t_2 the '9' neuron in the output layer receives strong excitatory conductances from the teaching signal, while the other neurons receive strong inhibitory conductances (Figure 3C). This phase lasts from t_1 to t_2, and at t_2 another set of plateau potentials occurs in the hidden layer neurons. The result is that we have plateaus potentials in the hidden layer neurons for both the forward ($\boldsymbol{\alpha}^f = [\alpha_1^f, ..., \alpha_m^f]$) and target ($\boldsymbol{\alpha}^t = [\alpha_1^t, ..., \alpha_m^t]$) phases:

$$\alpha_i^f = \sigma(\int_{t_1-t_p}^{t_1} A_i(t)dt)$$
$$\alpha_i^t = \sigma(\int_{t_2-t_p}^{t_2} A_i(t)dt) \quad (3)$$

where t_p is the integration time for the plateau potential (see Materials and Methods). Note that since the plateau potentials are generated using the same sigmoid non-linearity that we use to calculate the rate-of-fire at the soma, they are analogous to the firing rates of the hidden layer neurons. This allows us to use the plateau potentials to define target firing rates for the neurons (see below). The network is simulated in near continuous-time (except that each plateau is considered to be instantaneous), and the intervals between plateaus are randomly sampled from an inverse Gaussian distribution (Figure 3D, top). As such, the specific amount of time that the network is presented with each image and teaching signal is stochastic, though usually somewhere between 50-60 ms of simulated time (Figure 3D, bottom). This stochasticity was not necessary, but it demonstrates that although the system operates in phases, the specific length of the phases is not important as long as they are sufficiently long to permit integration (see Lemma 1). In the data presented in this paper, all 60,000 images in the MNIST training set were presented to the network one at a time, and each exposure to the full set of images was considered an "epoch" of training. At the end of each epoch, the network's classification error rate on a separate set of 10,000 test images was assessed with a single forward phase (see Materials and Methods).

It is important to note that there are many aspects of this design that are not physiologically accurate. Most notably, stochastic generation of synchronized plateau potentials across a population is not an accurate reflection of how real pyramidal neurons operate, since apical calcium spikes are determined by a number of concrete physiological factors in individual cells, including back-propagating action potentials, spike-timing and inhibitory inputs (Larkum et al., 1999, 2007, 2009). However, we note that calcium spikes in the apical dendrites can be prevented from occurring via the activity of distal dendrite targeting inhibitory interneurons (Murayama et al., 2009), which can synchronize pyramidal activity (Hilscher et al., 2017). Furthermore, distal dendrite targeting neurons can themselves can be rapidly inhibited in response to temporally precise neuromodulatory inputs (Pi et al., 2013; Pfeffer et al., 2013; Karnani et al., 2016; Hangya et al., 2015; Brombas et al., 2014). Therefore, it is entirely plausible that neocortical micro-circuits would generate synchronized pyramidal plateaus at punctuated periods of time in response to disinhibition of the apical dendrites governed by neuromodulatory signals that determine "phases" of processing. Alternatively, oscillations in population activity could provide a mechanism for promoting alternating phases of processing and synaptic plasticity (Buzsáki and Draguhn, 2004). But, complete synchrony of plateaus in our hidden layer neurons is not actually critical to our algorithm—only the temporal relationship between the plateaus and the teaching signal is critical (see below). This relationship itself is arguably plausible given the role of neuromodulatory inputs in dis-inhibiting the distal dendrites of pyramidal neurons (Karnani et al., 2016; Brombas et al., 2014). Of course, we are engaged in a great deal of speculation here. But, the point is that our model utilizes anatomical and functional motifs that are analogous to what is observed in the neocortex. Importantly for the present study, the key issue is the use of segregated dendrites and distinct transmit and plateau modes to solve the credit assignment problem.

Credit assignment with segregated dendrites

To solve the credit assignment problem without using weight transport, we had to define a local error function for the hidden layer that somehow takes into account the impact that each hidden layer neuron has on the output at the final layer. In other words, credit assignment could only be achieved if we knew that changing a hidden layer synapse to reduce the local error would also help to reduce error at the output layer. To obtain this guarantee, we defined local targets for the output and the hidden layer, i.e. desired firing rates for both the output layer neurons and the hidden layer neurons. Learning is then

a process of changing the synaptic connections to achieve these target firing rates across the network. Importantly, we defined our hidden layer targets in a similar manner to Lee et al. (2015). These hidden layer targets coordinate with the output layer by incorporating the feedback information contained in the plateau potentials. In this way, credit assignment is accomplished through mechanisms that are spatially local to the hidden layer neurons.

Specifically, to create the output and hidden layer targets, we use the average rates-of-fire and average voltages during different phases. We define:

$$\begin{aligned} \overline{\boldsymbol{X}}^y &= [\overline{X_1}^y, ..., \overline{X_k}^y] \\ \overline{X_i}^f &= \int_{t_1-t_p}^{t_1} X_i(t)dt \\ \overline{X_i}^t &= \int_{t_2-t_p}^{t_2} X_i(t)dt \end{aligned} \tag{4}$$

where $y \in \{f, t\}$, $k \in \{\ell, m, n\}$ and $X_i(t)$ can be a voltage variable (e.g. $C_i(t)$ or $U_i(t)$) or a rate-of-fire variable (e.g. $\lambda_i^C(t)$ or $\lambda_i^U(t)$). Using the averages over different phases, we then define the target rates-of-fire for the output layer, $\hat{\boldsymbol{\lambda}}^U = [\hat{\lambda}_1^U, ..., \hat{\lambda}_n^U]$, to be the average rates-of-fire during the target phase:

$$\hat{\lambda}_i^U = \overline{\lambda_i^U}^t \tag{5}$$

For the hidden layer we define the target rates-of-fire, $\hat{\boldsymbol{\lambda}}^C = [\hat{\lambda}_1^C, ..., \hat{\lambda}_m^C]$, using the average rates-of-fire during the forward phase and the difference between the plateau potentials from the forward and transmit phase:

$$\hat{\lambda}_i^C = \overline{\lambda_i^C}^f + \alpha_i^t - \alpha_i^f \tag{6}$$

The goal of learning in the hidden layer is to change the synapses $\boldsymbol{W}^0$ to achieve these targets in response to the given inputs.

More generally, for both the output layer and the hidden layer we then define error functions, L^1 and L^0, respectively, based on the difference between the local targets and the activity of the neurons during the forward phase (Figure 4A):

$$\begin{aligned} L^1 &= ||\hat{\boldsymbol{\lambda}}^U - \lambda_{max}\sigma(\overline{\boldsymbol{U}}^f)||_2^2 \\ L^0 &= ||\hat{\boldsymbol{\lambda}}^C - \lambda_{max}\sigma(\overline{\boldsymbol{C}}^f)||_2^2 \end{aligned} \tag{7}$$

Note that the loss function for the hidden layer, L^0, will, on average, reduce to the difference $||\boldsymbol{\alpha}^t - \boldsymbol{\alpha}^f||_2^2$. This provides an intuitive reason for why these targets help with credit assignment. Specifically, a hidden layer neuron's error is small only when the output layer is providing similar feedback to it during the forward and target phases. Put another way, hidden layer neurons "know" that they are "doing well" when they receive the same feedback from the output layer regardless of whether or not the teaching signal is present. Thus, hidden layer neurons have access to some information about how they are doing in helping the output layer to achieve its targets.

More formally, it can be shown that the output and hidden layer error functions will, on average, agree with each other, i.e. if the hidden layer error is reduced then the output layer error is also reduced. Specifically, similar to the proof employed by Lee et al. (2015), it can be shown that if the output layer error is sufficiently small and the synaptic matrices $\boldsymbol{W}^1$ and $\boldsymbol{Y}$ meet some conditions, then:

$$||\hat{\boldsymbol{\lambda}}^U - \lambda_{max}\sigma(k_D \boldsymbol{W}^1 \hat{\boldsymbol{\lambda}}^C)||_2^2 < ||\hat{\boldsymbol{\lambda}}^U - \lambda_{max}\sigma(E[\overline{\boldsymbol{U}}^f])||_2^2 \tag{8}$$

where k_D is a conductance term and $E[\cdot]$ denotes the expected value. (See Theorem 1 for the proof and a concrete description of the conditions). In plain language, equation (8) says that when we consider the difference between the output target and the activity that the hidden target *would* have induced at the output layer, it is less than the difference between the output target and the expected value of the output layer activity during the forward phase. In other words, the output layer's error would have, on average, been *smaller* if the hidden layer had achieved its own target during the forward phase. Hence, if we reduce the hidden layer error, it should also reduce the output layer error, thereby providing a guarantee of appropriate credit assignment.

With these error functions, we update the weights in the network with local gradient descent:

$$\begin{aligned} \Delta \boldsymbol{W}^1 &\propto \frac{\partial L^1}{\partial \boldsymbol{W}^1} \\ \Delta \boldsymbol{W}^0 &\propto \frac{\partial L^0}{\partial \boldsymbol{W}^0} \end{aligned} \tag{9}$$

where $\Delta \boldsymbol{W}^i$ refers to the update term for weight matrix $\boldsymbol{W}^i$ (see Materials and Methods, equations (19), (20), (22) and (23) for details of the weight update procedures). Given the coordination implied by equation (8), as the hidden layer reduces its own local error with gradient descent, the output layer's error should also be reduced, i.e. hidden layer learning should imply output layer learning.

To test that we were successful in credit assignment for the hidden layer, and to provide empirical support for the proof, we compared the local error at the hidden layer to the output layer error across all of the image presentations to the network. We observed that, generally, whenever the hidden layer error was low, the output layer error was also low. For example, when we consider the errors for the set of '2' images presented to the network during the second epoch, there was a Pearson correlation coefficient between L^0 and L^1 of $r = 0.61$, which was much higher than what was observed for shuffled data, wherein output and hidden activities were randomly paired (Figure 4B). Furthermore, these correlations were observed across all epochs of training, with most correlation coefficients for the hidden and output errors falling between $r = 0.2$ - 0.6, which was, again, much higher than the correlations observed for shuffled data (Figure 4C).

Interestingly, the correlations between L^0 and L^1 were smaller on the first epoch of training. This suggests that the coordination between the layers may only come into full effect once the network has engaged in some learning. Therefore, we inspected whether the conditions on the synaptic matrices that are assumed in the proof were, in fact, being met. More precisely, the proof assumes that the feedforward and feedback synaptic matrices ($\boldsymbol{W}^1$ and $\boldsymbol{Y}$, respectively) produce forward and backward transformations between the output and hidden layer that are approximate inverses of each other (see Proof of Theorem 1). Since we begin learning with random matrices, this condition is almost definitely *not* met at the start of training. But, we found that the network learned to meet this condition. Inspection of $\boldsymbol{W}^1$ and $\boldsymbol{Y}$ showed that during the first epoch the forward and backwards functions became approximate inverses of each other (Figure 4, Supplement 1). This means that during the first few image presentations the network was actually *learning to do credit assignment*. This result is very similar to previous models examining feedback alignment (Lillicrap et al., 2016), and shows that the feedback alignment (Lillicrap et al., 2016) and difference target propagation (Lee et al., 2015) algorithms are intimately linked. Furthermore, our results suggest that very early development may involve a period of learning how to assign credit appropriately. Altogether, our model demonstrates that credit assignment using random feedback weights is a general principle that can be implemented using segregated dendrites.

Deep learning with segregated dendrites

Given our finding that the network was successfully assigning credit for the output error to the hidden layer neurons, we had reason to believe that our network with local weight-updates would exhibit deep

learning, i.e. an ability to take advantage of a multi-layer structure (Bengio and LeCun, 2007). To test this, we examined the effects of including hidden layers. If deep learning is indeed operational in the network, then the inclusion of hidden layers should improve the ability of the network to classify images.

We built three different versions of the network (Figure 5A). The first was a network that had no hidden layer, i.e. the input neurons projected directly to the output neurons. The second was the network illustrated in Figure 3A, with a single hidden layer. The third contained two hidden layers, with the output layer projecting directly back to both hidden layers. This direct projection allowed us to build our local targets for each hidden layer using the plateaus driven by the output layer, thereby avoiding a "backward pass" through the entire network as has been used in other models (Lillicrap et al., 2016; Lee et al., 2015; Liao et al., 2015). We trained each network on the 60,000 MNIST training images for 60 epochs, and recorded the percentage of images in the 10,000 test image set that were incorrectly classified. The network with no hidden layers rapidly learned to classify the images, but it also rapidly hit an asymptote at an average error rate of 8.3% (Figure 5B, gray line). In contrast, the network with one hidden layer did not exhibit a rapid convergence to an asymptote in its error rate. Instead, it continued to improve throughout all 60 epochs, achieving an average error rate of 4.1% by the 60^{th} epoch (Figure 5B, blue line). Similar results were obtained when we loosened the synchrony constraints and instead allowed each hidden layer neuron to engage in plateau potentials at different times (Figure 5, Supplement 1). This demonstrates that strict synchrony in the plateau potentials is not required. But, our target definitions do require two different plateau potentials separated by the teaching signal input, which mandates some temporal control of plateau potentials in the system.

Interestingly, we found that the addition of a second hidden layer further improved learning. The network with two hidden layers learned more rapidly than the network with one hidden layer and achieved an average error rate of 3.2% on the test images by the 60^{th} epoch, also without hitting a clear asymptote in learning (Figure 5B, red line). However, it should be noted that additional hidden layers beyond two did not significantly improve the error rate (data not shown), which suggests that our particular algorithm could not be used to construct very deep networks as is. Nonetheless, our network was clearly able to take advantage of multi-layer architectures to improve its learning, which is the key feature of deep learning (Bengio and LeCun, 2007; LeCun et al., 2015).

Another key feature of deep learning is the ability to generate representations in the higher layers of a network that capture task-relevant information while discarding sensory details (LeCun et al., 2015; Mnih et al., 2015). To examine whether our network exhibited this type of abstraction, we used the t-Distributed Stochastic Neighbor Embedding algorithm (t-SNE). The t-SNE algorithm reduces the dimensionality of data while preserving local structure and non-linear manifolds that exist in high-dimensional space, thereby allowing accurate visualization of the structure of high-dimensional data (Maaten and Hinton, 2008). We applied t-SNE to the activity patterns at each layer of the two hidden layer network for all of the images in the test set after 60 epochs of training. At the input level, there was already some clustering of images based on their categories. However, the clusters were quite messy, with different categories showing outliers, several clusters, or merged clusters (Figure 5C, bottom). For example, the '2' digits in the input layer exhibited two distinct clusters separated by a cluster of '7's: one cluster contained '2's with a loop and one contained '2's without a loop. Similarly, there were two distinct clusters of '4's and '9's that were very close to each other, with one pair for digits on a pronounced slant and one for straight digits (Figure 5C, bottom, example images). Thus, although there is built-in structure to the categories of the MNIST dataset, there are a number of low-level features that do not respect category boundaries. In contrast, at the first hidden layer, the activity patterns were much cleaner, with far fewer outliers and split/merged clusters (Figure 5C, middle). For example, the two separate '2' digit clusters were much closer to each other and were now only separated by a very small cluster of '7's. Likewise, the '9' and '4' clusters were now distinct and no longer split based on the slant of the digit. Interestingly, when we examined the activity patterns at the second hidden layer the categories were even better segregated with only a bit of splitting or merging of category clusters (Figure 5C, top). Therefore, the network had learned to develop representations in the hidden layers wherein the categories were very distinct and low-level features unrelated to the categories were largely ignored. This abstract representation is likely to be key to the improved error rate in the two hidden layer network.

Altogether, our data demonstrates that our network with segregated dendritic compartments can engage in deep learning.

Coordinated local learning mimics backpropagation of error

The backpropagation of error algorithm (Rumelhart et al., 1986) is still the primary learning algorithm used for deep supervised learning in artificial neural networks (LeCun et al., 2015). Previous work has shown that learning with random feedback weights can actually match the synaptic weight updates specified by the backpropagation algorithm after a few epochs of training (Lillicrap et al., 2016). This fascinating observation suggests that deep learning with random feedback weights is not so much a different algorithm than backpropagation of error, but rather, networks with random feedback connections learn to approximate credit assignment as it is done in backpropagation (Lillicrap et al., 2016). Hence, we were curious as to whether or not our network was, in fact, learning to approximate the synaptic weight updates prescribed by backpropagation. To test this, we trained our one hidden layer network as before, but now, in addition to calculating the vector of hidden layer synaptic weight updates specified by our local learning rule ($\Delta \boldsymbol{W}^0$ in equation (9)), we also calculated the vector of hidden layer synaptic weight updates that would be specified by non-locally backpropagating the error from the output layer, ($\Delta \boldsymbol{W}^0_{BP}$). We then calculated the angle between these two alternative weight updates. In a very high-dimensional space, any two independent vectors will be roughly orthogonal to each other (i.e. $\Delta \boldsymbol{W}^0 \angle \Delta \boldsymbol{W}^0_{BP} \approx 90°$). If the two synaptic weight update vectors are *not* orthogonal to each other (i.e. $\Delta \boldsymbol{W}^0 \angle \Delta \boldsymbol{W}^0_{BP} < 90°$), then it suggests that the two algorithms are specifying similar weight updates.

As in previous work (Lillicrap et al., 2016), we found that the initial weight updates for our network were orthogonal to the updates specified by backpropagation. But, as the network learned the angle dropped to approximately 65°, before rising again slightly to roughly 70° (Figure 6A, blue line). This suggests that our network was learning to develop local weight updates in the hidden layer that were in rough agreement with the updates that explicit backpropagation would produce. However, this drop in orthogonality was still much less than that observed in non-spiking artificial neural networks learning with random feedback weights, which show a drop to below 45° (Lillicrap et al., 2016). We suspected that the higher angle between the weight updates that we observed may have been because we were using spikes to communicate the feedback from the upper layer, which could introduce both noise and bias in the estimates of the output layer activity. To test this, we also examined the weight updates that our algorithm would produce if we propagated the spike rates of the output layer neurons, $\boldsymbol{\lambda}^U(t)$, back directly through the random feedback weights, $\boldsymbol{Y}$. In this scenario, we observed a much sharper drop in the $\Delta \boldsymbol{W}^0 \angle \Delta \boldsymbol{W}^0_{BP}$ angle, which reduced to roughly 35° before rising again to 40° (Figure 6A, red line). These results show that, in principle, our algorithm is learning to approximate the backpropagation algorithm, though with some drop in accuracy introduced by the use of spikes to propagate output layer activities to the hidden layer.

To further examine how our local learning algorithm compared to backpropagation we compared the low-level features that the two algorithms learned. To do this, we trained the one hidden layer network with both our algorithm and backpropagation. We then examined the receptive fields (i.e. the synaptic weights) produced by both algorithms in the hidden layer synapses ($\boldsymbol{W}^0$) after 60 epochs of training. The two algorithms produced qualitatively similar receptive fields (Figure 6C). Both produced receptive fields with clear, high-contrast features for detecting particular strokes or shapes. To quantify the similarity, we conducted pair-wise correlation calculations for the receptive fields produced by the two algorithms and identified the maximum correlation pairs for each. Compared to shuffled versions of the receptive fields, there was a very high level of maximum correlation (Figure 6B), showing that the receptive fields were indeed quite similar. Thus, the data demonstrate that our learning algorithm using random feedback weights into segregated dendrites can in fact come to approximate the backpropagation of error algorithm.

Conditions on feedback weights

Once we had convinced ourselves that our learning algorithm was, in fact, producing deep learning similar to that produced by backpropagation of error, we wanted to examine some of the constraints on learning. First, we wanted to explore the structure of the feedback weights. In our initial simulations we used non-sparse, random (i.e. normally distributed) feedback weights. We were interested in whether learning could still work with sparse weights, given that neocortical connectivity is sparse. As well, we wondered whether symmetric weights would *improve* learning, which would be expected given previous findings (Lillicrap et al., 2016; Lee et al., 2015; Liao et al., 2015). To explore these questions, we trained our one hidden layer network using both sparse feedback weights (only 20% non-zero values) and symmetric weights ($\boldsymbol{Y} = \boldsymbol{W}^{1^T}$) (Figure 7A,C). We found that learning actually *improved* slightly with sparse weights (Figure 7B, red line), achieving an average error rate of 3.7% by the 60^{th} epoch, compared to the average 4.1% error rate achieved with fully random weights. But, this result appeared to depend on the magnitude of the sparse weights. To compensate for the loss of 80% of the weights we initially increased the sparse synaptic weight magnitudes by a factor of 5. However, when we did not re-scale the sparse weights learning was actually *worse* (Figure 7, Supplement 1). This suggests that sparse feedback provides a signal that is sufficient for credit assignment, but only if it is of appropriate magnitude.

Similar to sparse feedback weights, symmetric feedback weights also improved learning, leading to a rapid decrease in the test error and an error rate of 3.6% by the 60^{th} epoch (Figure 7D, red line). However, when we added noise to the symmetric weights this advantage was eliminated and learning was, in fact, slightly impaired (Figure 7D, blue line). At first, this was a very surprising result: given that learning works with random feedback weights, why would it not work with symmetric weights with noise? However, when we considered our previous finding that during the first epoch the feedforward weights, $\boldsymbol{W}^1$, learn to match the inverse of the feedback function (Figure 4, Supplement 1) a possible answer becomes clear. In the case of symmetric feedback weights the synaptic matrix $\boldsymbol{Y}$ is changing as $\boldsymbol{W}^1$ changes. This works fine when $\boldsymbol{Y}$ is set to $\boldsymbol{W}^{1^T}$, since that artificially forces weight alignment. But, if the feedback weights are set to $\boldsymbol{W}^{1^T}$ plus noise, then the system can never align the weights appropriately, since $\boldsymbol{Y}$ is now a moving target. This would imply that any implementation of feedback learning must either be very effective (to achieve the right feedback) or very slow (to allow the feedforward weights to adapt).

Learning with partial apical attenuation

Another constraint that we wished to examine was whether total segregation of the apical inputs as we had done was necessary, given that real pyramidal neurons only show an attenuation of distal apical inputs to the soma (Larkum et al., 1999). To examine this, we re-ran our two hidden layer network, but now, we allowed the apical dendritic voltage to influence the somatic voltage by setting $g_A = 0.05$. This value gave us twelve times more attenuation than the attenuation from the basal compartments (since $g_B = 0.6$). This difference in the levels of attenuation is in-line with experimental data (Larkum et al., 1999, 2009) (Figure 8A). When we compared the learning in this scenario to the scenario with total apical segregation, we observed very little difference in the error rates on the test set (Figure 8B, gray and red lines). Importantly, though, we found that if we decreased the apical attenuation to the same level as the basal compartment ($g_A = g_B = 0.6$) then the learning was significantly impaired (Figure 8B, blue line). This demonstrates that although total apical attenuation is not necessary, partial segregation of the apical compartment from the soma is necessary. This result makes sense given that our local targets for the hidden layer neurons incorporate a term that is supposed to reflect the response of the output neurons to the feedforward sensory information ($\boldsymbol{\alpha}^f$). Without some sort of separation of feedforward and feedback information, as is assumed in other models of deep learning (Lillicrap et al., 2016; Lee et al., 2015), this feedback signal would get corrupted by recurrent dynamics in the network. Our data show that electrontonically segregated dendrites is one potential way to achieve the required separation between feedforward and feedback information.

Discussion

Deep learning has radically altered the field of AI, demonstrating that parallel distributed processing across multiple layers can produce human/animal-level capabilities in image classification, pattern recognition and reinforcement learning (Hinton et al., 2006; LeCun et al., 2015; Mnih et al., 2015; Silver et al., 2016; Krizhevsky et al., 2012; He et al., 2015). Deep learning was motivated by analogies to the real brain (LeCun et al., 2015; Cox and Dean, 2014), so it is tantalizing that recent studies have shown that deep neural networks develop representations that strongly resemble the representations observed in the mammalian neocortex (Khaligh-Razavi and Kriegeskorte, 2014; Yamins and DiCarlo, 2016; Cadieu et al., 2014; Kubilius et al., 2016). In fact, deep learning models can match cortical representations even more so than some models that explicitly attempt to mimic the real brain (Khaligh-Razavi and Kriegeskorte, 2014). Hence, at a phenomenological level, it appears that deep learning, defined as multilayer cost function reduction with appropriate credit assignment, may be key to the remarkable computational prowess of the mammalian brain (Marblestone et al., 2016). However, the lack of biologically feasible mechanisms for credit assignment in deep learning algorithms, most notably backpropagation of error (Rumelhart et al., 1986), has left neuroscientists with a mystery. How can the real brain solve the credit assignment problem (Figure 1)? Here, we expanded on an idea that previous authors have explored (Körding and König, 2001; Spratling, 2002; Spratling and Johnson, 2006) and demonstrated that segregating the feedback and feedforward inputs to neurons, much as the real neocortex does (Larkum et al., 1999, 2007, 2009), can enable the construction of local targets to assign credit appropriately to hidden layer neurons (Figure 2). With this formulation, we showed that we could use segregated dendritic compartments to coordinate learning across layers (Figure 3 and Figure 4). This enabled our network to take advantage of multiple layers to develop representations of hand-written digits in hidden layers that enabled better levels of classification accuracy on the MNIST dataset than could be achieved with a single layer (Figure 5). Furthermore, we found that our algorithm actually approximated the weight updates that would be prescribed by backpropagation, and produced similar low-level feature detectors (Figure 6). As well, we showed that our basic framework works with sparse feedback connections (Figure 7) and more realistic, partial apical attenuation (Figure 8). Therefore, our work demonstrates that deep learning is possible in a biologically feasible framework, provided that feedforward and feedback signals are sufficiently segregated in different dendrites.

Perhaps the most biologically unrealistic component of backpropagation is the use of non-local "weight-transport" (Figure 1B) (Grossberg, 1987). Our model builds on recent neural networks research demonstrating that weight transport is not, in fact, required for credit assignment (Lillicrap et al., 2016; Liao et al., 2015; Lee et al., 2015). By demonstrating that the credit assignment problem is solvable using feedback to generate spatially local weight updates, these studies have provided a major breakthrough in our understanding of how deep learning could work in the real brain. However, this previous research involved an implicit assumption of separate feedforward and feedback pathways (Figure 2A). Although this is a possibility, there is currently no evidence for separate feedback pathways to guide learning in the neocortex. In our study, we obtained separate feedforward and feedback information using electrotonically segregated dendrites (Figure 2B), in analogy to neocortical pyramidal neurons (Larkum et al., 1999). This segregation allowed us to perform credit assignment using direct feedback pathways between layers (Figures 3-5). Moreover, we found that even with partial attenuation of the conductance from the apical dendrites to the soma our network could engage in deep learning (Figure 8). It should be recognized, though, that although our learning algorithm achieved deep learning with spatially local update rules, we had to assume some temporal non-locality. Our credit assignment system depended on the difference between the target and forward plateaus, $\boldsymbol{\alpha}^t - \boldsymbol{\alpha}^f$, which occurred at different times (roughly a 50-60 ms gap). Although this is not an ideal solution, this small degree of temporal non-locality for synaptic update rules is in-line with known biological mechanisms, such as slowly decaying calcium transients or synaptic tags (Redondo and Morris, 2011). Hence, our model exhibited deep learning using only local information contained within the cells.

In this work we adopted a similar strategy to the one taken by Lee et al.'s (2015) difference target propagation algorithm, wherein the feedback from higher layers is used to construct local activity targets

at the hidden layers. One of the reasons that we adopted this strategy is that it is appealing to think that feedback from upper layers may not simply be providing a signal for plasticity, but also a modulatory signal to push the hidden layer neurons towards a "better" activity pattern in *real-time*. This sort of top-down modulation could be used by the brain to improve sensory processing in different contexts and engage in inference (Bengio et al., 2015). Indeed, framing cortico-cortical feedback as a mechanism to modulate incoming sensory activity is a more common way of viewing feedback signals in the neocortex (Larkum, 2013; Gilbert and Li, 2013; Zhang et al., 2014; Fiser et al., 2016). In light of this, it is interesting to note that distal apical inputs in somatosensory cortex can help animals perform sensory discrimination tasks (Takahashi et al., 2016; Manita et al., 2015). However, in our model, we did not actually implement a system that altered the hidden layer activity to make sensory calculations—we simply used the feedback signals to drive learning. In-line with this view of top-down feedback, two recent papers have found evidence that cortical feedback can indeed guide feedforward sensory plasticity (Thompson et al., 2016; Yamada et al., 2017). Yet, there is no reason that feedback signals cannot provide both top-down modulation and a signal for learning (Spratling, 2002). In this respect, a potential future advance on our model would be to implement a system wherein the feedback actively "nudges" the hidden layers towards appropriate activity patterns in order to guide learning while also shaping perception. This proposal is reminiscent of the approach taken in previous computational models (Urbanczik and Senn, 2014; Spratling and Johnson, 2006; Körding and König, 2001). Future research could study how top-down modulation and a signal for credit assignment can be combined in deep learning models.

It is important to note that there are many aspects of our model that are not biologically realistic. These include (but are not limited to) synaptic inputs that are not conductance based, synaptic weights that can switch from positive to negative (and *vice-versa*), and plateau potentials that are considered instantaneous and are not an accurate reflection of calcium spike dynamics in real pyramidal neurons (Larkum et al., 1999, 2009). However, the intention with this study was not to achieve total biological realism, but rather to demonstrate that the sort of dendritic segregation that neocortical pyramidal neurons exhibit can subserve credit assignment. Although we found that deep learning can still work with realistic levels of passive conductance from the apical compartment to the somatic compartment (Figure 8B, red line) (Larkum et al., 1999), we also found that learning was impaired if the basal and apical compartments had equal conductance to the soma (Figure 8B, blue line). This is interesting, because it suggests that the unique physiology of neocortical pyramidal neurons may in fact reflect nature's solution to deep learning. Perhaps the relegation of feedback from higher-order brain regions to the electrotonically distant apical dendrites (Budd, 1998; Spratling, 2002), and the presence of active calcium spike mechanisms in the apical shaft (Larkum et al., 2009), are mechanisms for coordinating local synaptic weight updates. If this is correct, then the inhibitory interneurons that target apical dendrites and limit active communication to the soma (Murayama et al., 2009) may be used by the neocortex to control learning. Although this is speculative, it is worth noting that current evidence supports the idea that neuromodulatory inputs carrying temporally precise salience information (Hangya et al., 2015) can shut off interneurons to disinhibit the distal apical dendrites (Pi et al., 2013; Karnani et al., 2016; Pfeffer et al., 2013; Brombas et al., 2014), and presumably, promote apical communication to the soma. Recent work suggests that the specific patterns of interneuron inhibition on the apical dendrites are very spatially precise and differentially timed to motor behaviours (Muñoz et al., 2017), which suggests that there may well be coordinated physiological mechanisms for determining when cortico-cortical feedback is transmitted to the soma. Future research should examine whether these inhibitory and neuromodulatory mechanisms do, in fact, open the window for plasticity in the basal dendrites of pyramidal neurons, as our model, and some recent experimental work (Bittner et al., 2015), would predict.

An additional issue that should be recognized is that the error rates which our network achieved were by no means as low as can be achieved with artificial neural networks, nor at human levels of performance (LeCun et al., 1998; Li et al., 2016). As well, our algorithm was not able to take advantage of very deep structures (beyond two hidden layers, the error rate did not improve). In contrast, increasing the depth of networks trained with backpropagation can lead to performance improvements (Li et al., 2016). But, these observations do not mean that our network was not engaged in deep learning. First, it is interesting to note that although the backpropagation algorithm is several decades old (Rumelhart et al., 1986), it

was long considered to be useless for training networks with more than one or two hidden layers (Bengio and LeCun, 2007). Indeed, it was only the use of layer-by-layer training that initially led to the realization that deeper networks can achieve excellent performance (Hinton et al., 2006). Since then, both the use of very large datasets (with millions of examples), and additional modifications to the backpropagation algorithm, have been key to making backpropagation work well on deeper networks (Sutskever et al., 2013; LeCun et al., 2015). Future studies could examine how our algorithm could incorporate current techniques used in machine learning to work better on deeper architectures. Second, we stress that our network was not designed to match the state-of-the-art in machine learning, nor human capabilities. To test our basic hypothesis (and to run our leaky-integration and spiking simulations in a reasonable amount of time) we kept the network small, we stopped training before it reached its asymptote, and we did not implement any add-ons to the learning to improve the error rates, such as convolution and pooling layers, initialization tricks, mini-batch training, drop-out, momentum or RMSProp (Sutskever et al., 2013; Tieleman and Hinton, 2012; Srivastava et al., 2014). Indeed, it would be quite surprising if a relatively vanilla, small network like ours could come close to matching current performance benchmarks in machine learning. Third, although our network was able to take advantage of multiple layers to improve the error rate, there may be a variety of reasons that ever increasing depth didn't improve performance significantly. For example, our use of direct connections from the output layer to the hidden layers may have impaired the network's ability to coordinate synaptic updates between *hidden* layers. As well, given our finding that the use of spikes produced weight updates that were less well-aligned to backpropagation (Figure 6A) it is possible that deeper architectures require mechanisms to overcome the inherent noisiness of spikes.

One aspect of our model that we did not develop was the potential for learning at the feedback synapses. Although we used random synaptic weights for feedback, we also demonstrated that our model actually learns to meet the mathematical conditions required for credit assignment, namely the feedforward weights come to approximate the inverse of the feedback (Figure 4, Supplement 1). This suggests that it would be beneficial to develop a synaptic weight update rule for the feedback synapses that made this aspect of the learning better. Indeed, Lee et al. (2015) implemented an "inverse loss function" for their feedback synapses which promoted the development of feedforward and feedback functions that were roughly inverses of each other, leading to the emergence of auto-encoder functions in their network. In light of this, it is interesting to note that there is evidence for unique, "reverse" spike-timing-dependent synaptic plasticity rules in the distal apical dendrites of pyramidal neurons (Sjöström and Häusser, 2006; Letzkus et al., 2006), which have been shown to produce symmetric feedback weights and auto-encoder functions in artificial spiking networks (Burbank and Kreiman, 2012; Burbank, 2015). Thus, it is possible that early in development the neocortex actually learns cortico-cortical feedback connections that help it to assign credit for later learning. Our work suggests that any experimental evidence showing that feedback connections learn to approximate the inverse of feedforward connections could be considered as evidence for deep learning in the neocortex.

In summary, deep learning has had a huge impact on AI, but, to date, its impact on neuroscience has been limited. Nonetheless, given a number of findings in neurophysiology and modeling (Yamins and DiCarlo, 2016), there is growing interest in understanding how deep learning may actually be achieved by the real brain (Marblestone et al., 2016). Our results show that by moving away from point neurons, and shifting towards multi-compartment neurons that segregate feedforward and feedback signals, the credit assignment problem can be solved and deep learning can be achieved. Perhaps the dendritic anatomy of neocortical pyramidal neurons is important for nature's own deep learning algorithm.

Materials and Methods

Code for the model can be obtained from a GitHub repository (https://github.com/jordan-g/Segregated-Dendrite-Deep-Learning) (Guergiuev, 2017). For notational simplicity, we describe our model in the case of a network with only one hidden layer. We describe how this is extended to a network with multiple layers at the end of this section. As well, at the end of this section in Table 1 we provide a table listing

the parameter values we used for all of the simulations presented in this paper.

Neuronal dynamics

The network described here consists of an input layer with ℓ neurons, a hidden layer with m neurons, and an output layer with n neurons. Neurons in the input layer are simple Poisson spiking neurons whose rate-of-fire is determined by the intensity of image pixels (ranging from 0 - λ_{max}). Neurons in the hidden layer are modeled using three functional compartments—basal dendrites with voltages $\boldsymbol{B}(t) = [B_1(t), B_2(t), ..., B_m(t)]$, apical dendrites with voltages $\boldsymbol{A}(t) = [A_1(t), A_2(t), ..., A_m(t)]$, and somata with voltages $\boldsymbol{C}(t) = [C_1(t), C_2(t), ..., C_m(t)]$. Feedforward inputs from the input layer and feedback inputs from the output layer arrive at basal and apical synapses, respectively. At basal synapses, presynaptic spikes are translated into post-synaptic potentials $\boldsymbol{PSP^B}(t) = [PSP_1^B(t), PSP_2^B(t), ..., PSP_\ell^B(t)]$ given by:

$$PSP_j^B(t) = \sum_{s \in X_j} \kappa(t - s) \tag{10}$$

where X_j is the set of pre-synaptic spike times at synapses with presynaptic input neuron j. $\kappa(t) = (e^{-t/\tau_L} - e^{-t/\tau_s})\Theta(t)/(\tau_L - \tau_s)$ is the response kernel, where τ_s and τ_L are short and long time constants, and Θ is the Heaviside step function. The post-synaptic potentials at apical synapses, $\boldsymbol{PSP^A}(t) = [PSP_1^A(t), PSP_2^A(t), ..., PSP_n^A(t)]$, are modeled in the same manner. The basal and apical dendritic potentials for neuron i are then given by weighted sums of the post-synaptic potentials at either its basal or apical synapses:

$$\begin{aligned} B_i(t) &= \sum_{j=1}^{\ell} W_{ij}^0 PSP_j^B(t) + b_i^0 \\ A_i(t) &= \sum_{j=1}^{n} Y_{ij} PSP_j^A(t) \end{aligned} \tag{11}$$

where $\boldsymbol{b}^0 = [b_1^0, b_2^0, ..., b_m^0]$ are bias terms, $\boldsymbol{W}^0$ is the $m \times \ell$ matrix of feedforward weights for neurons in the hidden layer, and $\boldsymbol{Y}$ is the $m \times n$ matrix of their feedback weights. The somatic voltage for neuron i evolves with leak as:

$$\frac{dC_i(t)}{dt} = -g_L C_i(t) + g_B(B_i(t) - C_i(t)) + g_A(A_i(t) - C_i(t)) \tag{12}$$

where g_L is the leak conductance, g_B is the conductance from the basal dendrite to the soma, and g_A is the conductance from the apical dendrite to the soma. Equation (12) is identical to equation (1) in Results. Note that for simplicity's sake we are assuming a resting potential of 0 V and a membrane capacitance of 1 F, but these values are not important for the results.

The instantaneous firing rates of neurons in the hidden layer are given by $\boldsymbol{\lambda}^C(t) = [\lambda_1^C(t), \lambda_2^C(t), ..., \lambda_m^C(t)]$, where $\lambda_i^C(t)$ is the result of applying a nonlinearity, $\sigma(\cdot)$, to the somatic potential $C_i(t)$. We chose $\sigma(\cdot)$ to be a simple sigmoidal function, such that:

$$\lambda_i^C(t) = \lambda_{max}\sigma(C_i(t)) = \lambda_{max}\frac{1}{1 + e^{-C_i(t)}} \tag{13}$$

Here, λ_{max} is the maximum possible rate-of-fire for the neurons, which we set to 200 Hz. Note that equation (13) is identical to equation (2) in Results. Spikes are then generated using Poisson processes with these firing rates. We note that although the maximum rate was 200 Hz, the neurons rarely achieved

anything close to this rate, and the average rate of fire in the neurons during our simulations was 24 Hz, in-line with neocortical firing rates (Steriade et al., 2001).

Units in the output layer are modeled using only two compartments, dendrites with voltages $\boldsymbol{V}(t) = [V_1(t), V_2(t), ..., V_n(t)]$ and somata with voltages $\boldsymbol{U}(t) = [U_1(t), U_2(t), ..., U_n(t)]$. $V_i(t)$ is given by:

$$V_i(t) = \sum_{j=1}^{m} W_{ij}^1 PSP_j^V(t) + b_i^1 \tag{14}$$

where $\boldsymbol{PSP}^V(t) = [PSP_1^V(t), PSP_2^V(t), ..., PSP_m^V(t)]$ are the post-synaptic potentials at synapses that receive feedforward input from the hidden layer, and are calculated in the manner described by equation (10). $U_i(t)$ evolves as:

$$\frac{dU_i(t)}{dt} = -g_L U_i(t) + g_D(V_i(t) - U_i(t)) + I_i(t) \tag{15}$$

where g_L is the leak conductance, g_D is the conductance from the dendrite to the soma, and $\boldsymbol{I}(t) = [I_1(t), I_2(t), ..., I_n(t)]$ are somatic currents that can drive output neurons toward a desired somatic voltage. For neuron i, I_i is given by:

$$I_i(t) = g_{E_i}(t)(E_E - U_i(t)) + g_{I_i}(t)(E_I - U_i(t)) \tag{16}$$

where $\boldsymbol{g_E}(t) = [g_{E_1}(t), g_{E_2}(t), ..., g_{E_n}(t)]$ and $\boldsymbol{g_I}(t) = [g_{I_1}(t), g_{I_2}(t), ..., g_{I_n}(t)]$ are time-varying excitatory & inhibitory nudging conductances, and E_E and E_I are the excitatory and inhibitory reversal potentials. In our simulations, we set $E_E = 8$ V and $E_I = -8$ V. During the target phase only, we set $g_{I_i} = 1$ and $g_{E_i} = 0$ for all units i whose output should be minimal, and $g_{E_i} = 1$ and $g_{I_i} = 0$ for the unit whose output should be maximal. In this way, all units other than the "target" unit are silenced, while the "target" unit receives a strong excitatory drive. In the forward phase, $\boldsymbol{I}(t)$ is set to 0. The Poisson spike rates $\boldsymbol{\lambda}^U(t) = [\lambda_1^U(t), \lambda_2^U(t), ..., \lambda_n^U(t)]$ are calculated as in equation (13).

Plateau potentials

At the end of the forward and target phases, we calculate plateau potentials $\boldsymbol{\alpha}^f = [\alpha_1^f, \alpha_2^f, ..., \alpha_m^f]$ and $\boldsymbol{\alpha}^t = [\alpha_1^t, \alpha_2^t, ..., \alpha_m^t]$ for apical dendrites of hidden layer neurons, where α_i^f and α_i^t are given by:

$$\begin{aligned} \alpha_i^f &= \sigma(\int_{t_1 - t_p}^{t_1} A_i(t)dt) \\ \alpha_i^t &= \sigma(\int_{t_2 - t_p}^{t_2} A_i(t)dt) \end{aligned} \tag{17}$$

where t_1 and t_2 are the end times of the forward and target phases, respectively, and $t_p = t_{i+1} - (t_i + \Delta t_s)$ is the integration time for the plateau potential. $\Delta t_s = 30$ ms is the settling time for the voltages. Note that equation (17) is identical to equation (3) in Results. These plateau potentials are used by hidden layer neurons to update their basal weights.

Weight updates

All feedforward synaptic weights are updated at the end of each target phase. Output layer units update their synaptic weights $\boldsymbol{W}^1$ in order to minimize the loss function L^1 given in equation (7). All average voltages are calculated after a delay Δt_s from the start of a phase, which allows for the network to

reach a steady state before averaging begins. In practice this means that the average somatic voltage for output layer neuron i in the forward phase, $\overline{U_i}^f$, has the property

$$\overline{U_i}^f \approx k_D \overline{V_i}^f = k_D\big(\sum_{j=1}^{m} W_{ij}^1 \overline{PSP_j^V}^f + b_i^1\big) \tag{18}$$

where $k_D = g_D/(g_L + g_D)$. Thus,

$$\begin{aligned} \frac{\partial L^1}{\partial \boldsymbol{W}^1} &\approx -k_D \lambda_{max}(\hat{\boldsymbol{\lambda}}^U - \lambda_{max}\sigma(\overline{\boldsymbol{U}}^f))\sigma'(\overline{\boldsymbol{U}}^f) \circ \overline{\boldsymbol{PSP}^{\boldsymbol{V}}}^f \\ \frac{\partial L^1}{\partial \boldsymbol{b}^1} &\approx -k_D \lambda_{max}(\hat{\boldsymbol{\lambda}}^U - \lambda_{max}\sigma(\overline{\boldsymbol{U}}^f))\sigma'(\overline{\boldsymbol{U}}^f) \end{aligned} \tag{19}$$

The dendrites in the output layer use this approximation of the gradient in order to update their weights using gradient descent:

$$\begin{aligned} \boldsymbol{W}^1 &\to \boldsymbol{W}^1 - \eta^1 P^1 \frac{\partial L^1}{\partial \boldsymbol{W}^1} \\ \boldsymbol{b}^1 &\to \boldsymbol{b}^1 - \eta^1 P^1 \frac{\partial L^1}{\partial \boldsymbol{b}^1} \end{aligned} \tag{20}$$

where η^1 is a learning rate constant, and P^1 is a scaling factor used to normalize the scale of the rate-of-fire function.

In the hidden layer, basal dendrites update their synaptic weights $\boldsymbol{W}^0$ by minimizing the loss function L^0 given in equation (7). We define the target rates-of-fire $\hat{\boldsymbol{\lambda}}^C = [\hat{\lambda}_1^C, \hat{\lambda}_2^C, ..., \hat{\lambda}_m^C]$ such that

$$\hat{\lambda}_i^C = \overline{\lambda_i^C}^f + \alpha_i^t - \alpha_i^f \tag{21}$$

where $\boldsymbol{\alpha}^f = [\alpha_1^f, \alpha_2^f, ..., \alpha_m^f]$ and $\boldsymbol{\alpha}^t = [\alpha_1^t, \alpha_2^t, ..., \alpha_m^t]$ are forward and target phase plateau potentials given in equation (17). Note that equation (21) is identical to equation (6) in Results. These hidden layer target firing rates are similar to the targets used in difference target propagation (Lee et al., 2015). Using the fact that $\overline{\lambda_i^C}^f \approx \lambda_{max}\sigma(\overline{C_i}^f)$, we can say that $\hat{\lambda}_i^C - \lambda_{max}\sigma(\overline{C_i}^f) \approx \alpha_i^t - \alpha_i^f$, and hence:

$$\begin{aligned} \frac{\partial L^0}{\partial \boldsymbol{W}^0} &\approx -k_B(\boldsymbol{\alpha}^t - \boldsymbol{\alpha}^f)\lambda_{max}\sigma'(\overline{\boldsymbol{C}}^f) \circ \overline{\boldsymbol{PSP}^{\boldsymbol{B}}}^f \\ \frac{\partial L^0}{\partial \boldsymbol{b}^0} &\approx -k_B(\boldsymbol{\alpha}^t - \boldsymbol{\alpha}^f)\lambda_{max}\sigma'(\overline{\boldsymbol{C}}^f) \end{aligned} \tag{22}$$

where $k_B = g_B/(g_L + g_B + g_A)$. Basal weights are updated in order to descend this gradient:

$$\begin{aligned} \boldsymbol{W}^0 &\to \boldsymbol{W}^0 - \eta^0 P^0 \frac{\partial L^0}{\partial \boldsymbol{W}^0} \\ \boldsymbol{b}^0 &\to \boldsymbol{b}^0 - \eta^0 P^0 \frac{\partial L^0}{\partial \boldsymbol{b}^0} \end{aligned} \tag{23}$$

Importantly, this update rule is fully local for the hidden layer neurons. It consists essentially of three terms, (1) the difference in the plateau potentials for the target and forward phases ($\boldsymbol{\alpha}^t - \boldsymbol{\alpha}^f$), (2) the derivative of the spike rate function ($\lambda_{max}\sigma'(\overline{\boldsymbol{C}}^f)$), and (3) the postsynaptic potentials ($\overline{\boldsymbol{PSP}^{\boldsymbol{B}}}^f$). All three of these terms are values that a real neuron could theoretically calculate using some combination of molecular synaptic tags, calcium currents, and back-propagating action potentials.

Multiple hidden layers

In order to extend our algorithm to deeper networks with multiple hidden layers, our model incorporates direct synaptic connections from the output layer to each hidden layer. Thus, each hidden layer receives feedback from the output layer through its own separate set of fixed, random weights. For example, in a network with two hidden layers, both layers receive the feedback from the output layer at their apical dendrites through backward weights $\boldsymbol{Y}^0$ and $\boldsymbol{Y}^1$. The local targets at each layer are then given by:

$$\hat{\boldsymbol{\lambda}}^U = \overline{\boldsymbol{\lambda}^U}^t \tag{24}$$

$$\hat{\boldsymbol{\lambda}}^{C^1} = \overline{\boldsymbol{\lambda}^{C^1}}^t + \boldsymbol{\alpha}^{1\,t} - \boldsymbol{\alpha}^{1\,f} \tag{25}$$

$$\hat{\boldsymbol{\lambda}}^{C^0} = \overline{\boldsymbol{\lambda}^{C^0}}^t + \boldsymbol{\alpha}^{0\,t} - \boldsymbol{\alpha}^{0\,f} \tag{26}$$

where the superscripts ^{0} and ^{1} denote the first and second hidden layers, respectively.
The local loss functions at each layer are:

$$\begin{aligned} L^2 &= ||\hat{\boldsymbol{\lambda}}^U - \lambda_{max}\sigma(\overline{\boldsymbol{U}}^f)||_2^2 \\ L^1 &= ||\hat{\boldsymbol{\lambda}}^{C^1} - \lambda_{max}\sigma(\overline{\boldsymbol{C}^1}^f)||_2^2 \\ L^0 &= ||\hat{\boldsymbol{\lambda}}^{C^0} - \lambda_{max}\sigma(\overline{\boldsymbol{C}^0}^f)||_2^2 \end{aligned} \tag{27}$$

where L^2 is the loss at the output layer. The learning rules used by the hidden layers in this scenario are the same as in the case with one hidden layer.

Learning rate optimization

For each of the three network sizes that we present in this paper, a grid search was performed in order to find good learning rates. We set the learning rate for each layer by stepping through the range $[0.1, 0.3]$ with a step size of 0.02. For each combination of learning rates, a neural network was trained for one epoch on the 60, 000 training examples, after which the network was tested on 10,000 test images. The learning rates that gave the best performance on the test set after an epoch of training were used as a basis for a second grid search around these learning rates that used a smaller step size of 0.01. From this, the learning rates that gave the best test performance after 20 epochs were chosen as our learning rates for that network size.

In all of our simulations, we used a learning rate of 0.19 for a network with no hidden layers, learning rates of 0.21 (output and hidden) for a network with one hidden layer, and learning rates of 0.23 (hidden layers) and 0.12 (output layer) for a network with two hidden layers. All networks with one hidden layer had 500 hidden layer neurons, and all networks with two hidden layers had 500 neurons in the first hidden layer and 100 neurons in the second hidden layer.

Training paradigm

For all simulations described in this paper, the neural networks were trained on classifying handwritten digits using the MNIST database of 28 pixel $\times$ 28 pixel images. Initial feedforward and feedback weights were chosen randomly from a uniform distribution over a range that was calculated to produce voltages in the dendrites between -6 - 12 V.

Prior to training, we tested a network's initial performance on a set of 10,000 test examples. This set of images was shuffled at the beginning of testing, and each example was shown to the network in sequence. Each input image was encoded into Poisson spiking activity of the 784 input neurons representing each pixel of the image. The firing rate of an input neuron was proportional to the brightness of the pixel that it represents (with spike rates between [0 - λ_{max}]. The spiking activity of each of the

784 input neurons was received by the neurons in the first hidden layer. For each test image, the network underwent only a forward phase. At the end of this phase, the network's classification of the input image was given by the neuron in the output layer with the greatest somatic potential (and therefore the greatest spike rate). The network's classification was compared to the target classification. After classifying all 10,000 testing examples, the network's classification error was given by the percentage of examples that it did not classify correctly.

Following the initial test, training of the neural network was done in an on-line fashion. All 60,000 training images were randomly shuffled at the start of each training epoch. The network was then shown each training image in sequence, undergoing a forward phase ending with a plateau potential, and a target phase ending with another plateau potential. All feedforward weights were then updated at the end of the target phase. At the end of the epoch (after all 60,000 images were shown to the network), the network was again tested on the 10,000 test examples. The network was trained for up to 60 epochs.

Simulation details

For each training example, a minimum length of 50 ms was used for each of the forward and target phases. The lengths of the forward and target training phases were determined by adding their minimum length to an extra length term, which was chosen randomly from a Wald distribution with a mean of 2 ms and scale factor of 1. During testing, a fixed length of 500 ms was used for the forward transmit phase. Average forward and target phase voltages were calculated after a settle duration of $\Delta t_s = 30$ ms from the start of the phase.

For simulations with randomly sampled plateau potential times (Figure 5, Supplement 1), the time at which each neuron's plateau potential occurred was randomly sampled from a folded normal distribution ($\mu = 0, \sigma^2 = 3$) that was truncated ($\max = 5$) such that plateau potentials occurred between 0 ms and 5 ms before the start of the next phase. In this scenario, the average apical voltage in the last 30 ms was averaged in the calculation of the plateau potential for a particular neuron.

The time-step used for simulations was $dt = 1$ ms. At each time-step, the network's state was updated bottom-to-top beginning with the first hidden layer and ending with the output layer. For each layer, dendritic potentials were updated, followed by somatic potentials, and finally their spiking activity. Table 1 lists the simulation parameters and the values that were used in the figures presented.

All code was written using the Python programming language version 2.7 (RRID: SCR_008394) with the NumPy (RRID: SCR_008633) and SciPy (RRID: SCR_008058) libraries. The code is open source and is freely available at https://github.com/jordan-g/Segregated-Dendrite-Deep-Learning (Guergiuev, 2017). The data used to train the network was from the Mixed National Institute of Standards and Technology (MNIST) database, which is a modification of the original database from the National Institute of Standards and Technology (RRID: SCR_006440) (LeCun et al., 1998). The MNIST database can be found at http://yann.lecun.com/exdb/mnist/. Some of the simulations were run on the SciNet High-Performance Computing platform (Loken et al., 2010).

Proofs

Theorem for loss function coordination

The targets that we selected for the hidden layer (see equation (6)) were based on the targets used in Lee et al. (2015). The authors of that paper provided a proof showing that their hidden layer targets guaranteed that learning in one layer helped reduce the error in the next layer. However, there were a number of differences between our network and theirs, such as the use of spiking neurons, voltages, different compartments, etc. Here, we modify the original Lee et al. (2015) proof slightly to prove Theorem 1 (which is the formal statement of equation (8)).

One important thing to note is that the theorem given here utilizes a target for the hidden layer that is slightly different than the one defined in equation (6). However, the target defined in equation (6) is a

numerical approximation of the target given in Theorem 1. After the proof of we describe exactly how these approximations relate to the targets given here.

Theorem 1. *Consider a neural network with one hidden layer and an output layer. Let* $\hat{\boldsymbol{\lambda}}^{C^*} = \overline{\boldsymbol{\lambda}^{C}}^{f} + \sigma(\boldsymbol{Y}\overline{\boldsymbol{\lambda}^{U}}^{t}) - \sigma(\boldsymbol{Y}\lambda_{max}\sigma(E[\overline{\boldsymbol{U}}^{f}]))$ *be the target firing rates for neurons in the hidden layer, where* $\sigma(\cdot)$ *is a differentiable function. Assume that* $\overline{\boldsymbol{U}}^{f} \approx k_D\overline{\boldsymbol{V}}^{f}$. *Let* $\hat{\boldsymbol{\lambda}}^{U} = \overline{\boldsymbol{\lambda}^{U}}^{t}$ *be the target firing rates for the output layer. Also, for notational simplicity, let* $\beta(\boldsymbol{x}) \equiv \lambda_{max}\sigma(k_D\boldsymbol{W}^1\boldsymbol{x})$ *and* $\gamma(\boldsymbol{x}) \equiv \sigma(Y\boldsymbol{x})$. *Theorem 1 states that if* $\hat{\boldsymbol{\lambda}}^{U} - \lambda_{max}\sigma(E[\overline{\boldsymbol{U}}^{f}])$ *is sufficiently small, and the Jacobian matrices* J_β *and* J_γ *satisfy the condition that the largest eigenvalue of* $(I - J_\beta J_\gamma)^T(I - J_\beta J_\gamma)$ *is less than* 1, *then*

$$||\hat{\boldsymbol{\lambda}}^{U} - \lambda_{max}\sigma(k_D\boldsymbol{W}^1\hat{\boldsymbol{\lambda}}^{C^*})||_2^2 < ||\hat{\boldsymbol{\lambda}}^{U} - \lambda_{max}\sigma(E[\overline{\boldsymbol{U}}^{f}])||_2^2$$

We note again that the proof for this theorem is essentially a modification of the proof provided in Lee et al. (2015) that incorporates our Lemma 1 to take into account the expected value of $\overline{\boldsymbol{PSP}^{\boldsymbol{V}}}^{f}$, *given that spikes in the network are generated with non-stationary Poisson processes.*

Proof.

$$\begin{aligned}\hat{\boldsymbol{\lambda}}^{U} - \lambda_{max}\sigma(k_D\boldsymbol{W}^1\hat{\boldsymbol{\lambda}}^{C^*}) &\equiv \hat{\boldsymbol{\lambda}}^{U} - \beta(\hat{\boldsymbol{\lambda}}^{C^*}) \\ &= \hat{\boldsymbol{\lambda}}^{U} - \beta(\overline{\boldsymbol{\lambda}^{C}}^{f} + \gamma(\overline{\boldsymbol{\lambda}^{U}}^{t}) - \gamma(\lambda_{max}\sigma(E[\overline{\boldsymbol{U}}^{f}])))\end{aligned}$$

Lemma 1 shows that $\lambda_{max}\sigma(E[\overline{\boldsymbol{U}}^{f}]) = \lambda_{max}\sigma(E[k_D\boldsymbol{W}^1\overline{\boldsymbol{PSP}^{\boldsymbol{V}}}^{f}]) \approx \lambda_{max}\sigma(k_D\boldsymbol{W}^1\overline{\boldsymbol{\lambda}^{C}}^{f})$ given a sufficiently large averaging time window. Assume that $\lambda_{max}\sigma(E[\overline{\boldsymbol{U}}^{f}]) = \lambda_{max}\sigma(k_D\boldsymbol{W}^1\overline{\boldsymbol{\lambda}^{C}}^{f}) \equiv \beta(\overline{\boldsymbol{\lambda}^{C}}^{f})$. Then,

$$\hat{\boldsymbol{\lambda}}^{U} - \beta(\hat{\boldsymbol{\lambda}}^{C^*}) = \hat{\boldsymbol{\lambda}}^{U} - \beta(\overline{\boldsymbol{\lambda}^{C}}^{f} + \gamma(\overline{\boldsymbol{\lambda}^{U}}^{t}) - \gamma(\beta(\overline{\boldsymbol{\lambda}^{C}}^{f})))$$

Let $\boldsymbol{e} = \overline{\boldsymbol{\lambda}^{U}}^{t} - \beta(\overline{\boldsymbol{\lambda}^{C}}^{f})$. Applying Taylor's theorem,

$$\hat{\boldsymbol{\lambda}}^{U} - \beta(\hat{\boldsymbol{\lambda}}^{C^*}) = \hat{\boldsymbol{\lambda}}^{U} - \beta(\overline{\boldsymbol{\lambda}^{C}}^{f} + J_\gamma\boldsymbol{e} + \boldsymbol{o}(||\boldsymbol{e}||_2))$$

where $\boldsymbol{o}(||\boldsymbol{e}||_2)$ is the remainder term that satisfies $\lim_{\boldsymbol{e}\to 0}\boldsymbol{o}(||\boldsymbol{e}||_2)/||\boldsymbol{e}||_2 = 0$. Applying Taylor's theorem again,

$$\begin{aligned}\hat{\boldsymbol{\lambda}}^{U} - \beta(\hat{\boldsymbol{\lambda}}^{C^*}) &= \hat{\boldsymbol{\lambda}}^{U} - \beta(\overline{\boldsymbol{\lambda}^{C}}^{f}) - J_\beta(J_\gamma\boldsymbol{e} + \boldsymbol{o}(||\boldsymbol{e}||_2)) \\ &\quad - \boldsymbol{o}(||(J_\gamma\boldsymbol{e} + \boldsymbol{o}(||\boldsymbol{e}||_2)||_2) \\ &= \hat{\boldsymbol{\lambda}}^{U} - \beta(\overline{\boldsymbol{\lambda}^{C}}^{f}) + J_\beta J_\gamma\boldsymbol{e} - \boldsymbol{o}(||\boldsymbol{e}||_2) \\ &= (I - J_\beta J_\gamma)\boldsymbol{e} - \boldsymbol{o}(||\boldsymbol{e}||_2)\end{aligned}$$

Then,

$$\begin{aligned}||\hat{\boldsymbol{\lambda}}^{U} - \beta(\hat{\boldsymbol{\lambda}}^{C^*})||_2^2 &= ((I - J_\beta J_\gamma)\boldsymbol{e} - \boldsymbol{o}(||\boldsymbol{e}||_2))^T((I - J_\beta J_\gamma)\boldsymbol{e} - \boldsymbol{o}(||\boldsymbol{e}||_2)) \\ &= \boldsymbol{e}^T(I - J_\beta J_\gamma)^T(I - J_\beta J_\gamma)\boldsymbol{e} - \boldsymbol{o}(||\boldsymbol{e}||_2)^T(I - J_\beta J_\gamma)\boldsymbol{e} \\ &\quad - \boldsymbol{e}^T(I - J_\beta J_\gamma)^T\boldsymbol{o}(||\boldsymbol{e}||_2) + \boldsymbol{o}(||\boldsymbol{e}||_2)^T\boldsymbol{o}(||\boldsymbol{e}||_2) \\ &= \boldsymbol{e}^T(I - J_\beta J_\gamma)^T(I - J_\beta J_\gamma)\boldsymbol{e} + \boldsymbol{o}(||\boldsymbol{e}||_2^2) \\ &\leq \mu||\boldsymbol{e}||_2^2 + |o(||\boldsymbol{e}||_2^2)|\end{aligned}$$

where μ is the largest eigenvalue of $(I-J_\beta J_\gamma)^T(I-J_\beta J_\gamma)$. If $\boldsymbol{e}$ is sufficiently small so that $|o(||\boldsymbol{e}||_2^2))| < (1-\mu)||\boldsymbol{e}||_2^2$, then

$$||\hat{\boldsymbol{\lambda}}^U - \lambda_{max}\sigma(k_D \boldsymbol{W}^1 \hat{\boldsymbol{\lambda}}^{C^*})||_2^2 \leq ||\boldsymbol{e}||_2^2 = ||\hat{\boldsymbol{\lambda}}^U - \lambda_{max}\sigma(E[\overline{\boldsymbol{U}}^f])||_2^2$$

□

Note that the last step requires that μ, the largest eigenvalue of $(I - J_\beta J_\gamma)^T(I - J_\beta J_\gamma)$, is below 1. Clearly, we do not actually have any guarantee of meeting this condition. However, our results show that even though the feedback weights are random and fixed, the feedforward weights actually learn to meet this condition during the first epoch of training (Figure 4, Supplement 1).

Hidden layer targets

Theorem 1 shows that if we use a target $\hat{\boldsymbol{\lambda}}^{C^*} = \overline{\boldsymbol{\lambda}^C}^f + \sigma(\boldsymbol{Y}\overline{\boldsymbol{\lambda}^U}^t) - \sigma(\boldsymbol{Y}\lambda_{max}\sigma(k_D \boldsymbol{W}^1 \overline{\boldsymbol{\lambda}^C}^f))$ for the hidden layer, there is a guarantee that the hidden layer approaching this target will also push the upper layer closer to its target $\hat{\boldsymbol{\lambda}}^U$, if certain other conditions are met. Our specific choice of $\hat{\boldsymbol{\lambda}}^C$ defined in the Results (equation (6)) approximates this target rate vector using variables that are accessible to the hidden layer units.

Because neuronal units calculate averages after the network has reached a steady state, $\lambda_{max}\sigma(\overline{\boldsymbol{U}}^f) \approx \overline{\boldsymbol{\lambda}^U}^f$ and $\overline{\boldsymbol{A}}^f \approx \boldsymbol{Y}\overline{\boldsymbol{\lambda}^U}^f$. Using Lemma 1 and equation (4), $E[\overline{\boldsymbol{U}}^f] \approx k_D \boldsymbol{W}^1 \overline{\boldsymbol{\lambda}^C}^f$. If we assume that $\overline{\boldsymbol{U}}^f \approx E[\overline{\boldsymbol{U}}^f]$, which is true on average, then:

$$\boldsymbol{\alpha}^f = \sigma(\overline{\boldsymbol{A}}^f) \approx \sigma(\boldsymbol{Y}\overline{\boldsymbol{\lambda}^U}^f) \approx \sigma(\boldsymbol{Y}\lambda_{max}\sigma(\overline{\boldsymbol{U}}^f)) \approx \sigma(\boldsymbol{Y}\lambda_{max}\sigma(k_D \boldsymbol{W}^1 \overline{\boldsymbol{\lambda}^C}^f)) \tag{28}$$

and:

$$\boldsymbol{\alpha}^t = \sigma(\overline{\boldsymbol{A}}^t) \approx \sigma(\boldsymbol{Y}\overline{\boldsymbol{\lambda}^U}^t) \tag{29}$$

Therefore, $\hat{\boldsymbol{\lambda}}^C \approx \hat{\boldsymbol{\lambda}}^{C^*}$.

Thus, our hidden layer targets ensure that our model employs a learning rule similar to difference target propagation that approximates the necessary conditions to guarantee error convergence.

Lemma for firing rates

Theorem 1 had to rely on the equivalence between the average spike rates of the neurons and the postsynaptic potentials. Here, we prove a lemma showing that this equivalence does indeed hold as long as the integration time is long enough relative to the synaptic time constants t_s and t_L.

Lemma 1. *Let X be a set of presynaptic spike times during the time interval $\Delta t = t_1 - t_0$, distributed according to an inhomogeneous Poisson process. Let $N = |X|$ denote the number of presynaptic spikes during this time window, and let $s_i \in X$ denote the i^{th} presynaptic spike, where $0 < i \leq N$. Finally, let $\lambda(t)$ denote the time-varying presynaptic firing rate (i.e. the time-varying mean of the Poisson process), and $PSP(t)$ be the postsynaptic potential at time t given by equation (10). Then, during the time window Δt, as long as $\Delta t \gg 2\tau_L^2\tau_s^2\overline{\lambda^2}/(\tau_L - \tau_s)^2(\tau_L + \tau_s)$,*

$$E[\overline{PSP(t)}] \approx \overline{\lambda}$$

Proof. The average of $PSP(t)$ over the time window Δt is

$$\begin{aligned}\overline{PSP} &= \frac{1}{\Delta t}\int_{t_0}^{t_1} PSP(t)dt \\ &= \frac{1}{\Delta t}\sum_{s\in X}\int_{t_0}^{t_1}\frac{e^{-(t-s)/\tau_L}-e^{-(t-s)/\tau_s}}{\tau_L-\tau_s}\Theta(t-s)dt\end{aligned}$$

Since $\Theta(t-s)=0$ for all $t<s$,

$$\begin{aligned}\overline{PSP} &= \frac{1}{\Delta t}\sum_{s\in X}\int_{s}^{t_1}\frac{e^{-(t-s)/\tau_L}-e^{-(t-s)/\tau_s}}{\tau_L-\tau_s}dt \\ &= \frac{1}{\Delta t}\left(N-\sum_{s\in X}\frac{\tau_L e^{-(t_1-s)/\tau_L}-\tau_s e^{-(t_1-s)/\tau_s}}{\tau_L-\tau_s}\right)\end{aligned}$$

The expected value of $\overline{PSP}$ with respect to X is given by

$$\begin{aligned}E_X[\overline{PSP}] &= E_X\left[\frac{1}{\Delta t}\left(N-\sum_{s\in X}\frac{\tau_L e^{-(t_1-s)/\tau_L}-\tau_s e^{-(t_1-s)/\tau_s}}{\tau_L-\tau_s}\right)\right] \\ &= \frac{E_X[N]}{\Delta t}-\frac{1}{\Delta t}E_X\left[\sum_{i=0}^{N}\left(\frac{\tau_L e^{-(t_1-s_i)/\tau_L}-\tau_s e^{-(t_1-s_i)/\tau_s}}{\tau_L-\tau_s}\right)\right]\end{aligned}$$

Since the presynaptic spikes are an inhomogeneous Poisson process with a rate λ, $E_X[N]=\int_{t_0}^{t_1}\lambda dt$. Thus,

$$\begin{aligned}E_X[\overline{PSP}] &= \frac{1}{\Delta t}\int_{t_0}^{t_1}\lambda dt-\frac{1}{\Delta t}E_X\left[\sum_{i=0}^{N}g(s_i)\right] \\ &= \overline{\lambda}-\frac{1}{\Delta t}E_X\left[\sum_{i=0}^{N}g(s_i)\right]\end{aligned}$$

where we let $g(s_i)\equiv(\tau_L e^{-(t_1-s_i)/\tau_L}-\tau_s e^{-(t_1-s_i)/\tau_s})/(\tau_L-\tau_s)$. Then, the law of total expectation gives

$$\begin{aligned}E_X\left[\sum_{i=0}^{N}g(s_i)\right] &= E_N\left[E_X\left[\sum_{i=0}^{N}g(s_i)\middle|N\right]\right] \\ &= \sum_{n=0}^{\infty}\left(E_X\left[\sum_{i=0}^{N}g(s_i)\middle|N=n\right]\cdot P(N=n)\right)\end{aligned}$$

Letting $f_{s_i}(s)$ denote $P(s_i=s)$, we have that

$$\begin{aligned}E_X\left[\sum_{i=0}^{N}g(s_i)\middle|N=n\right] &= \sum_{i=0}^{n}E_X[g(s_i)] \\ &= \sum_{i=0}^{n}\int_{t_0}^{t_1}g(s)f_{s_i}(s)ds\end{aligned}$$

Since Poisson spike times are independent, for an inhomogeneous Poisson process:

$$\begin{aligned} f_{s_i}(s) &= \frac{\lambda(s)}{\int_{t_0}^{t_1} \lambda(t)dt} \\ &= \frac{\lambda(s)}{\overline{\lambda}\Delta t} \end{aligned}$$

for all $s \in [t_0, t_1]$. Since Poisson spike times are independent, this is true for all i. Thus,

$$\begin{aligned} E_X\left[\sum_{i=0}^{N} g(s_i)\middle| N = n\right] &= \frac{1}{\overline{\lambda}\Delta t}\sum_{i=0}^{n}\int_{t_0}^{t_1} g(s)\lambda(s)ds \\ &= \frac{n}{\overline{\lambda}\Delta t}\int_{t_0}^{t_1} g(s)\lambda(s)ds \end{aligned}$$

Then,

$$\begin{aligned} E_X\left[\sum_{i=0}^{N} g(s_i)\right] &= \sum_{n=0}^{\infty}\left(\frac{n}{\overline{\lambda}\Delta t}\left(\int_{t_0}^{t_1} g(s)\lambda(s)ds\right)\cdot P(N=n)\right) \\ &= \frac{1}{\overline{\lambda}\Delta t}\left(\int_{t_0}^{t_1} g(s)\lambda(s)ds\right)\left(\sum_{n=0}^{\infty} n\cdot P(N=n)\right) \end{aligned}$$

Now, for an inhomogeneous Poisson process with time-varying rate $\lambda(t)$,

$$\begin{aligned} P(N=n) &= \frac{[\int_{t_0}^{t_1}\lambda(t)dt]^n e^{-\int_{t_0}^{t_1}\lambda(t)dt}}{n!} \\ &= \frac{[\overline{\lambda}\Delta t]^n e^{-(\overline{\lambda}\Delta t)}}{n!} \end{aligned}$$

Thus,

$$\begin{aligned} E_X\left[\sum_{i=0}^{N} g(s_i)\right] &= \frac{e^{-(\overline{\lambda}\Delta t)}}{\overline{\lambda}\Delta t}\left(\int_{t_0}^{t_1} g(s)\lambda(s)ds\right)\left(\sum_{n=0}^{\infty} n\frac{[\overline{\lambda}\Delta t]^n}{n!}\right) \\ &= \frac{e^{-(\overline{\lambda}\Delta t)}}{\overline{\lambda}\Delta t}\left(\int_{t_0}^{t_1} g(s)\lambda(s)ds\right)(\overline{\lambda}\Delta t)e^{\overline{\lambda}\Delta t} \\ &= \int_{t_0}^{t_1} g(s)\lambda(s)ds \end{aligned}$$

Then,

$$E_X[\overline{PSP}] = \overline{\lambda} - \frac{1}{\Delta t}\left(\int_{t_0}^{t_1} g(s)\lambda(s)ds\right)$$

The second term of this equation is always greater than or equal to 0, since $g(s) \geq 0$ and $\lambda(s) \geq 0$ for all s. Thus, $E_X[\overline{PSP}] \leq \overline{\lambda}$. As well, the Cauchy-Schwarz inequality states that

$$\int_{t_0}^{t_1} g(s)\lambda(s)ds \leq \sqrt{\int_{t_0}^{t_1} g(s)^2 ds}\sqrt{\int_{t_0}^{t_1} \lambda(s)^2 ds}$$
$$= \sqrt{\int_{t_0}^{t_1} g(s)^2 ds}\sqrt{\overline{\lambda^2}\Delta t}$$

where

$$\int_{t_0}^{t_1} g(s)^2 ds = \int_{t_0}^{t_1} \Big(\frac{\tau_L e^{-(t_1-s)/\tau_L} - \tau_s e^{-(t_1-s)/\tau_s}}{\tau_L - \tau_s}\Big)^2 ds$$
$$\leq \frac{1}{2(\tau_L - \tau_s)^2}\left(4\frac{\tau_L^2\tau_s^2}{\tau_L + \tau_s}\right)$$
$$= \frac{2\tau_L^2\tau_s^2}{(\tau_L - \tau_s)^2(\tau_L + \tau_s)}$$

Thus,

$$\int_{t_0}^{t_1} g(s)\lambda(s)ds \leq \sqrt{\frac{2\tau_L^2\tau_s^2}{(\tau_L - \tau_s)^2(\tau_L + \tau_s)}}\sqrt{\overline{\lambda^2}\Delta t}$$
$$= \sqrt{\Delta t}\sqrt{\frac{2\tau_L^2\tau_s^2\overline{\lambda^2}}{(\tau_L - \tau_s)^2(\tau_L + \tau_s)}}$$

Therefore,

$$E_X[\overline{PSP}] \geq \overline{\lambda} - \frac{1}{\Delta t}\sqrt{\Delta t}\sqrt{\frac{2\tau_L^2\tau_s^2\overline{\lambda^2}}{(\tau_L - \tau_s)^2(\tau_L + \tau_s)}}$$
$$= \overline{\lambda} - \sqrt{\frac{2\tau_L^2\tau_s^2\overline{\lambda^2}}{\Delta t(\tau_L - \tau_s)^2(\tau_L + \tau_s)}}$$

Then,

$$\overline{\lambda} - \sqrt{\frac{2\tau_L^2\tau_s^2\overline{\lambda^2}}{\Delta t(\tau_L - \tau_s)^2(\tau_L + \tau_s)}} \leq E_X[\overline{PSP}] \leq \overline{\lambda}$$

Thus, as long as $\Delta t \gg 2\tau_L^2\tau_s^2\overline{\lambda^2}/(\tau_L - \tau_s)^2(\tau_L + \tau_s)$, $E_X[\overline{PSP}] \approx \overline{\lambda}$. □

What this lemma says, effectively, is that the expected value of PSP is going to be roughly the average presynaptic rate of fire as long as the time over which the average is taken is sufficiently long in comparison to the postsynaptic time constants and the average rate-of-fire is sufficiently small. In our simulations, Δt is always greater than or equal to 50 ms, the average rate-of-fire is approximately 20 Hz, and our time constants τ_L and τ_s are 10 ms and 3 ms, respectively. Hence, in general:

$$2\tau_L^2\tau_s^2\overline{\lambda^2}/(\tau_L - \tau_s)^2(\tau_L + \tau_s) = 2(10)^2(3)^2(0.02)^2/(10-3)^2(10+3)$$
$$\approx 0.001$$
$$\ll 50$$

Thus, in the proof of Theorem 1, we assume $E_X[\overline{PSP}] = \overline{\lambda}$.

Parameter	Units	Value	Description
dt	ms	1	Time step resolution
λ_{max}	Hz	200	Maximum spike rate
τ_s	ms	3	Short synaptic time constant
τ_L	ms	10	Long synaptic time constant
Δt_s	ms	30	Settle duration for calculation of average voltages
g_B	S	0.6	Hidden layer conductance from basal dendrite to the soma
g_A	S	0, 0.05, 0.6	Hidden layer conductance from apical dendrite to the soma
g_D	S	0.6	Output layer conductance from dendrite to the soma
g_L	S	0.1	Leak conductance
P_0	–	$20/\lambda_{max}$	Hidden layer error signal scaling factor
P_1	–	$20/\lambda_{max}^2$	Output layer error signal scaling factor

Table 1. List of parameter values used in our simulations.

Acknowledgments

We would like to thank Douglas Tweed, João Sacramento, Walter Senn, and Yoshua Bengio for helpful discussions on this work. This research was supported by two grants to B.A.R.: a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2014-04947) and a Google Faculty Research Award. The authors declare no competing financial interests. Some simulations were performed on the gpc supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto.

References

Bengio, Y. and LeCun, Y. (2007). Scaling learning algorithms towards AI. *Large-scale kernel machines*, 34(5):1–41.

Bengio, Y., Lee, D.-H., Bornschein, J., and Lin, Z. (2015). Towards biologically plausible deep learning. *arXiv preprint arXiv:1502.04156*.

Bittner, K. C., Grienberger, C., Vaidya, S. P., Milstein, A. D., Macklin, J. J., Suh, J., Tonegawa, S., and Magee, J. C. (2015). Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons. *Nat Neurosci*, 18(8):1133–1142.

Brombas, A., Fletcher, L. N., and Williams, S. R. (2014). Activity-Dependent Modulation of Layer 1 Inhibitory Neocortical Circuits by Acetylcholine. *The Journal of Neuroscience*, 34(5):1932.

Budd, J. M. (1998). Extrastriate feedback to primary visual cortex in primates: a quantitative analysis of connectivity. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1400):1037–1044.

Burbank, K. S. (2015). Mirrored STDP Implements Autoencoder Learning in a Network of Spiking Neurons. *PLoS Comput Biol*, 11(12):e1004566.

Burbank, K. S. and Kreiman, G. (2012). Depression-Biased Reverse Plasticity Rule Is Required for Stable Learning at Top-Down Connections. *PLoS Comput Biol*, 8(3):e1002393.

Buzsáki, G. and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, 304(5679):1926–1929.

Cadieu, C. F., Hong, H., Yamins, D. L. K., Pinto, N., Ardila, D., Solomon, E. A., Majaj, N. J., and DiCarlo, J. J. (2014). Deep neural networks rival the representation of primate IT cortex for core visual object recognition. *PLoS Comput Biol*, 10(12):e1003963.

Cox, D. and Dean, T. (2014). Neural Networks and Neuroscience-Inspired Computer Vision. *Current Biology*, 24(18):R921–R929.

Crick, F. (1989). The recent excitement about neural networks. *Nature*, 337(6203):129–132.

Dan, Y. and Poo, M.-M. (2004). Spike timing-dependent plasticity of neural circuits. *Neuron*, 44(1):23–30.

Fiser, A., Mahringer, D., Oyibo, H. K., Petersen, A. V., Leinweber, M., and Keller, G. B. (2016). Experience-dependent spatial expectations in mouse visual cortex. *Nat Neurosci*, 19(12):1658–1664.

Gadagkar, V., Puzerey, P. A., Chen, R., Baird-Daniel, E., Farhang, A. R., and Goldberg, J. H. (2016). Dopamine neurons encode performance error in singing birds. *Science*, 354(6317):1278.

Gilbert, C. D. and Li, W. (2013). Top-down influences on visual processing. *Nat Rev Neurosci*, 14(5):350–363.

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

Guergiuev, J. (2017). Segregated-dendrite-deep-learning. Github. `https://github.com/jordan-g/Segregated-Dendrite-Deep-Learning`. 23f2c66.

Hangya, B., Ranade, S., Lorenc, M., and Kepecs, A. (2015). Central cholinergic neurons are rapidly recruited by reinforcement feedback. *Cell*, 162(5):1155–1168.

Harris, K. D. (2008). Stability of the fittest: organizing learning through retroaxonal signals. *Trends in Neurosciences*, 31(3):130–136.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1026–1034.

Hilscher, M. M., Leão, R. N., Edwards, S. J., Leão, K. E., and Kullander, K. (2017). Chrna2-Martinotti Cells Synchronize Layer 5 Type A Pyramidal Cells via Rebound Excitation. *PLOS Biology*, 15(2):e2001392.

Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554.

Karnani, M. M., Jackson, J., Ayzenshtat, I., Hamzehei Sichani, A., Manoocheri, K., Kim, S., and Yuste, R. (2016). Opening holes in the blanket of inhibition: Localized lateral disinhibition by VIP interneurons. *The Journal of Neuroscience*, 36(12):3471–3480.

Khaligh-Razavi, S.-M. and Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLoS Comput Biol*, 10(11):e1003915.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.

Kubilius, J., Bracci, S., and Op de Beeck, H. P. (2016). Deep neural networks as a computational model for human shape sensitivity. *PLoS Comput Biol*, 12(4):e1004896.

Körding, K. P. and König, P. (2001). Supervised and Unsupervised Learning with Two Sites of Synaptic Integration. *Journal of Computational Neuroscience*, 11(3):207–215.

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

Larkum, M. E., Nevian, T., Sandler, M., Polsky, A., and Schiller, J. (2009). Synaptic Integration in Tuft Dendrites of Layer 5 Pyramidal Neurons: A New Unifying Principle. *Science*, 325(5941):756–760.

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

Larkum, M. E., Zhu, J. J., and Sakmann, B. (1999). A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature*, 398(6725):338–341.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

Lee, D.-H., Zhang, S., Fischer, A., and Bengio, Y. (2015). Difference target propagation. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 498–515. Springer.

Leibo, J. Z., Liao, Q., Anselmi, F., Freiwald, W. A., and Poggio, T. (2017). View-Tolerant Face Recognition and Hebbian Learning Imply Mirror-Symmetric Neural Tuning to Head Orientation. *Current Biology*, 27(1):62–67.

Letzkus, J. J., Kampa, B. M., and Stuart, G. J. (2006). Learning rules for spike timing-dependent plasticity depend on dendritic synapse location. *Journal of Neuroscience*, 26(41):10420–10429.

Li, Y., Li, H., Xu, Y., Wang, J., and Zhang, Y. (2016). Very deep neural network for handwritten digit recognition. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 174–182. Springer.

Liao, Q., Leibo, J. Z., and Poggio, T. (2015). How Important is Weight Symmetry in Backpropagation? *arXiv preprint arXiv:1510.05067*.

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.

Loken, C., Gruner, D., Groer, L., Peltier, R., Bunn, N., Craig, M., Henriques, T., Dempsey, J., Yu, C.-H., Chen, J., Dursi, L. J., Chong, J., Northrup, S., Pinto, J., Knecht, N., and Zon, R. V. (2010). Scinet: Lessons learned from building a power-efficient top-20 system and data centre. *Journal of Physics: Conference Series*, 256(1):012026.

Maaten, L. v. d. and Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(Nov):2579–2605.

Malenka, R. C. and Bear, M. F. (2004). LTP and LTD. *Neuron*, 44(1):5–21.

Manita, S., Suzuki, T., Homma, C., Matsumoto, T., Odagawa, M., Yamada, K., Ota, K., Matsubara, C., Inutsuka, A., Sato, M., Ohkura, M., Yamanaka, A., Yanagawa, Y., Nakai, J., Hayashi, Y., Larkum, M., and Murayama, M. (2015). A top-down cortical circuit for accurate sensory perception. *Neuron*, 86(5):1304–1316.

Marblestone, A., Wayne, G., and Kording, K. (2016). Towards an integration of deep learning and neuroscience. *arXiv preprint arXiv:1606.03813*.

Martin, S. J., Grimwood, P. D., and Morris, R. G. M. (2000). Synaptic plasticity and memory: an evaluation of the hypothesis. *Annual Review of Neuroscience*, 23(1):649–711.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., and Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533.

Murayama, M., Perez-Garci, E., Nevian, T., Bock, T., Senn, W., and Larkum, M. E. (2009). Dendritic encoding of sensory stimuli controlled by deep cortical interneurons. *Nature*, 457(7233):1137–1141.

Muñoz, W., Tremblay, R., Levenstein, D., and Rudy, B. (2017). Layer-specific modulation of neocortical dendritic inhibition during active wakefulness. *Science*, 355(6328):954.

Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat Neurosci*, 16(8):1068–1076.

Pi, H.-J., Hangya, B., Kvitsiani, D., Sanders, J. I., Huang, Z. J., and Kepecs, A. (2013). Cortical interneurons that specialize in disinhibitory control. *Nature*, 503(7477):521–524.

Redondo, R. L. and Morris, R. G. M. (2011). Making memories last: the synaptic tagging and capture hypothesis. *Nat Rev Neurosci*, 12(1):17–30.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323:9.

Scellier, B. and Bengio, Y. (2016). Towards a biologically plausible backprop. *arXiv preprint arXiv:1602.05179*.

Schiller, J., Major, G., Koester, H. J., and Schiller, Y. (2000). NMDA spikes in basal dendrites of cortical pyramidal neurons. *Nature*, 404(6775):285–289.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., and Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489.

Sjöström, P. J. and Häusser, M. (2006). A Cooperative Switch Determines the Sign of Synaptic Plasticity in Distal Dendrites of Neocortical Pyramidal Neurons. *Neuron*, 51(2):227–238.

Spratling, M. W. (2002). Cortical Region Interactions and the Functional Role of Apical Dendrites. *Behavioral and Cognitive Neuroscience Reviews*, 1(3):219–228.

Spratling, M. W. and Johnson, M. H. (2006). A feedback model of perceptual learning and categorization. *Visual Cognition*, 13(2):129–165.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958.

Steriade, M., Timofeev, I., and Grenier, F. (2001). Natural Waking and Sleep States: A View From Inside Neocortical Neurons. *Journal of Neurophysiology*, 85(5):1969–1985.

Sutskever, I., Martens, J., Dahl, G. E., and Hinton, G. E. (2013). On the importance of initialization and momentum in deep learning. *ICML (3)*, 28:1139–1147.

Takahashi, N., Oertner, T. G., Hegemann, P., and Larkum, M. E. (2016). Active cortical dendrites modulate perception. *Science*, 354(6319):1587.

Thompson, A., Picard, N., Min, L., Fagiolini, M., and Chen, C. (2016). Cortical Feedback Regulates Feedforward Retinogeniculate Refinement. *Neuron*, 91(5):1021–1033.

Tieleman, T. and Hinton, G. (2012). Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. *COURSERA: Neural Networks for Machine Learning*, 4(2).

Urbanczik, R. and Senn, W. (2009). Reinforcement learning in populations of spiking neurons. *Nature Neuroscience*, 12(3):250.

Urbanczik, R. and Senn, W. (2014). Learning by the dendritic prediction of somatic spiking. *Neuron*, 81(3):521–528.

Yamada, Y., Bhaukaurally, K., Madarász, T. J., Pouget, A., Rodriguez, I., and Carleton, A. (2017). Context- and Output Layer-Dependent Long-Term Ensemble Plasticity in a Sensory Circuit. *Neuron*, 93(5):1198–1212.e5.

Yamins, D. L. K. and DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nat Neurosci*, 19(3):356–365.

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J. P., Chang, W.-C., Jenvay, S., Miyamichi, K., Luo, L., and Dan, Y. (2014). Long-range and local circuits for top-down modulation of visual cortex processing. *Science*, 345(6197):660–665.

Zylberberg, J., Murphy, J. T., and DeWeese, M. R. (2011). A Sparse Coding Model with Synaptically Local Plasticity and Spiking Neurons Can Account for the Diverse Shapes of V1 Simple Cell Receptive Fields. *PLOS Computational Biology*, 7(10):e1002250.

Figures

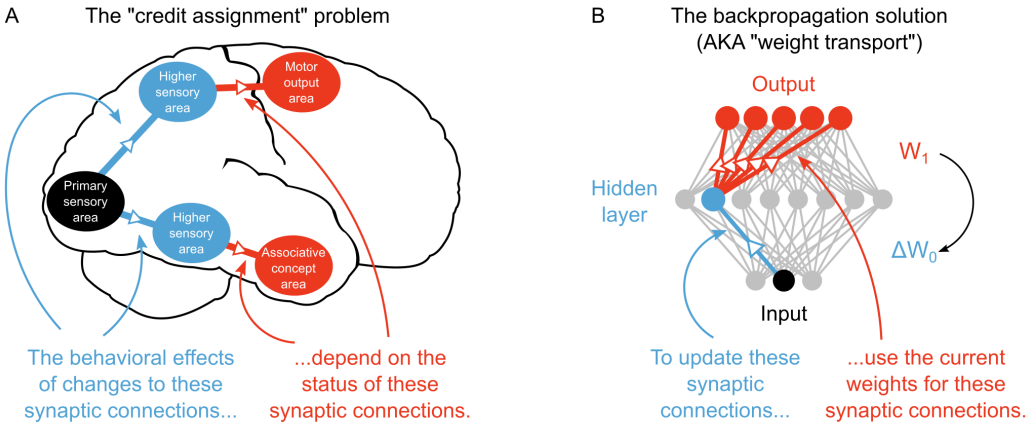

Figure 1. The credit assignment problem in multi-layer neural networks.

(**A**) Illustration of the credit assignment problem. In order to take full advantage of the multi-circuit architecture of the neocortex when learning, synapses in earlier processing stages (blue connections) must somehow receive "credit" for their impact on behavior or cognition. However, the credit due to any given synapse early in a processing pathway depends on the downstream synaptic connections that link the early pathway to later computations (red connections). (**B**) Illustration of weight transport in backpropagation. To solve the credit assignment problem, the backpropagation of error algorithm explicitly calculates the credit due to each synapse in the hidden layer by using the downstream synaptic weights when calculating the hidden layer weight changes. This solution works well in AI applications, but is unlikely to occur in the real brain.

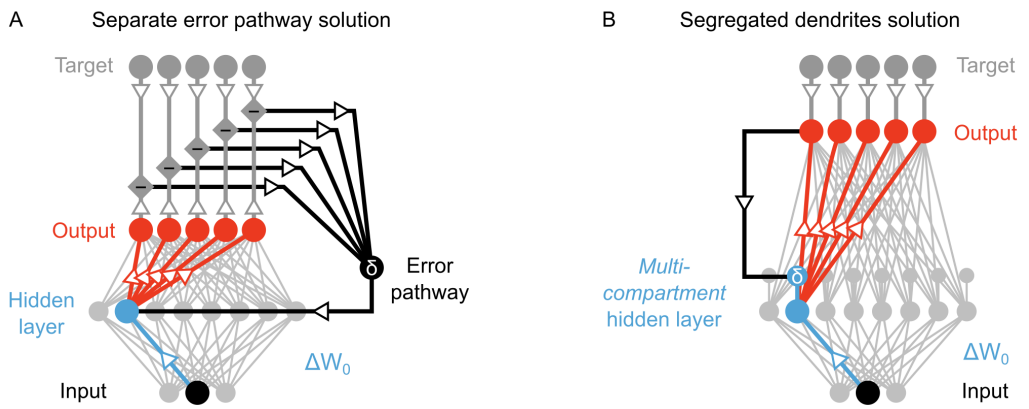

Figure 2. Potential solutions to credit assignment using top-down feedback.

(**A**) Illustration of the implicit feedback pathway used in previous models of deep learning. In order to assign credit, feedforward information must be integrated separately from any feedback signals used to calculate error for synaptic updates (the error is indicated here with δ). (**B**) Illustration of the segregated dendrites proposal. Rather than using a separate pathway to calculate error based on feedback, segregated dendritic compartments could receive feedback and calculate the error signals locally.

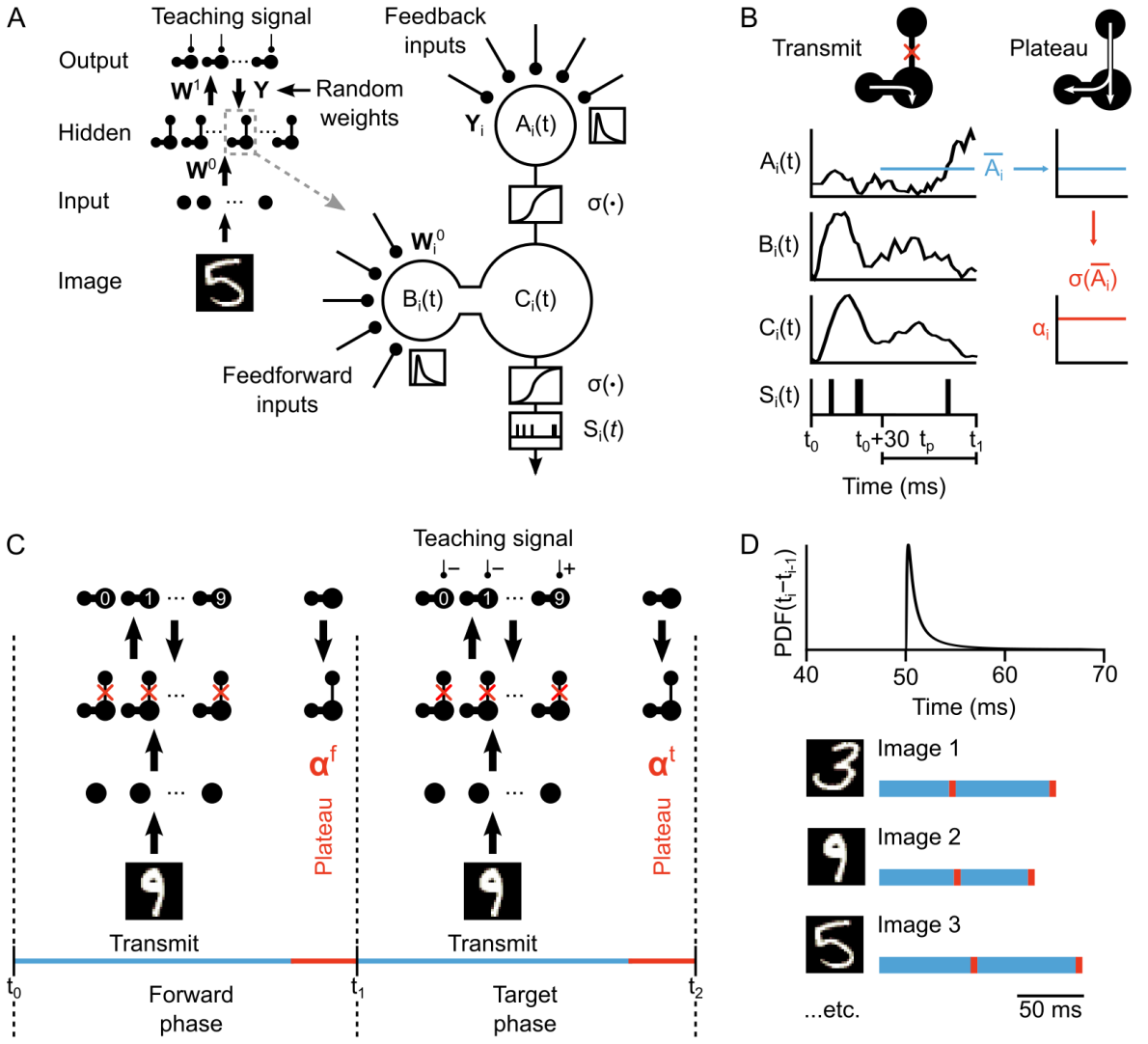

Figure 3. Illustration of a multi-compartment neural network model for deep learning.

(**A**) *Left*: Diagram of our network architecture. An input image is represented by the spiking of input units which propagate to the hidden layer through weights $\boldsymbol{W}^0$. Hidden layer activities arrive at the output layer through weights $\boldsymbol{W}^1$. Feedback from the output layer is sent back to the hidden layer through fixed, random weights $\boldsymbol{Y}$. *Right*: Illustration of unit i in the hidden layer. The unit has a basal dendrite, soma, and apical dendrite with membrane potentials B_i, C_i and A_i, respectively. The apical dendrite communicates to the soma predominantly using non-linear signals $\sigma(\cdot)$. The spiking output of the cell, $S_i(t)$, is a Poisson process with rate $\lambda_{max}\sigma(C_i(t))$. (**B**) Illustration of neuronal dynamics in the two modes of processing. In the transmit mode, the apical dendrite accumulates a measure of its average membrane potential from $t_1 - t_p$ to t_1. In the plateau mode, an apical plateau potential equal to the nonlinearity $\sigma(\cdot)$ applied to the average apical potential travels to the soma and basal dendrite. (**C**) Illustration of the sequence of network phases that occur for each training example. The network undergoes a forward phase (transmit & plateau) and a target phase (transmit & plateau) in sequence. In the target phase, a teaching signal is introduced to the output layer, shaping its activity. (**D**) Illustration of phase length sampling. For each training example, the lengths of forward & target phases are randomly drawn from a shifted inverse Gaussian distribution with a minimum of 50 ms.

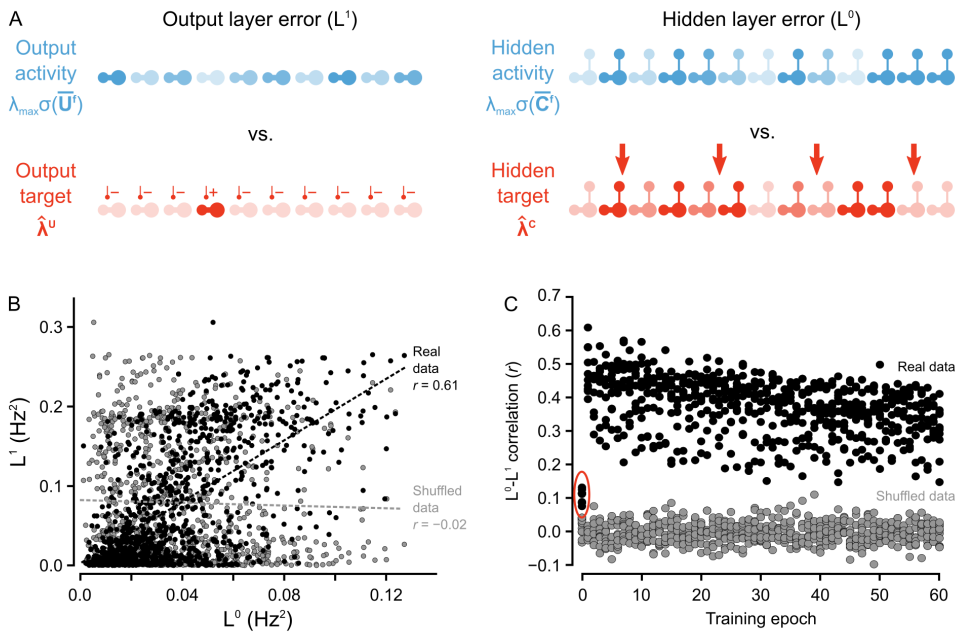

Figure 4. Co-ordinated errors between the output and hidden layers.

(**A**) Illustration of output error (L^1) and local hidden error (L^0). For a given test example shown to the network in a forward phase, the output layer error is defined as the squared norm of the difference between target firing rates $\hat{\boldsymbol{\lambda}}^U$ and the sigmoid function applied to the average somatic potentials $\overline{\boldsymbol{U}}^f$ across all output units. Hidden layer error is defined similarly, except the target is $\hat{\boldsymbol{\lambda}}^C$ (as defined in the text). (**B**) Plot of L^1 vs. L^0 for all of the '2' images after one epoch of training. There is a strong correlation between hidden layer error and output layer error (real data, black), as opposed to when output and hidden errors were randomly paired (shuffled data, gray). (**C**) Plot of correlation between hidden layer error and output layer error across training for each category of images (each dot represents one category). The correlation is significantly higher in the real data than the shuffled data throughout training. Note also that the correlation is much lower on the first epoch of training (red oval), suggesting that the conditions for credit assignment are still developing during the first epoch.

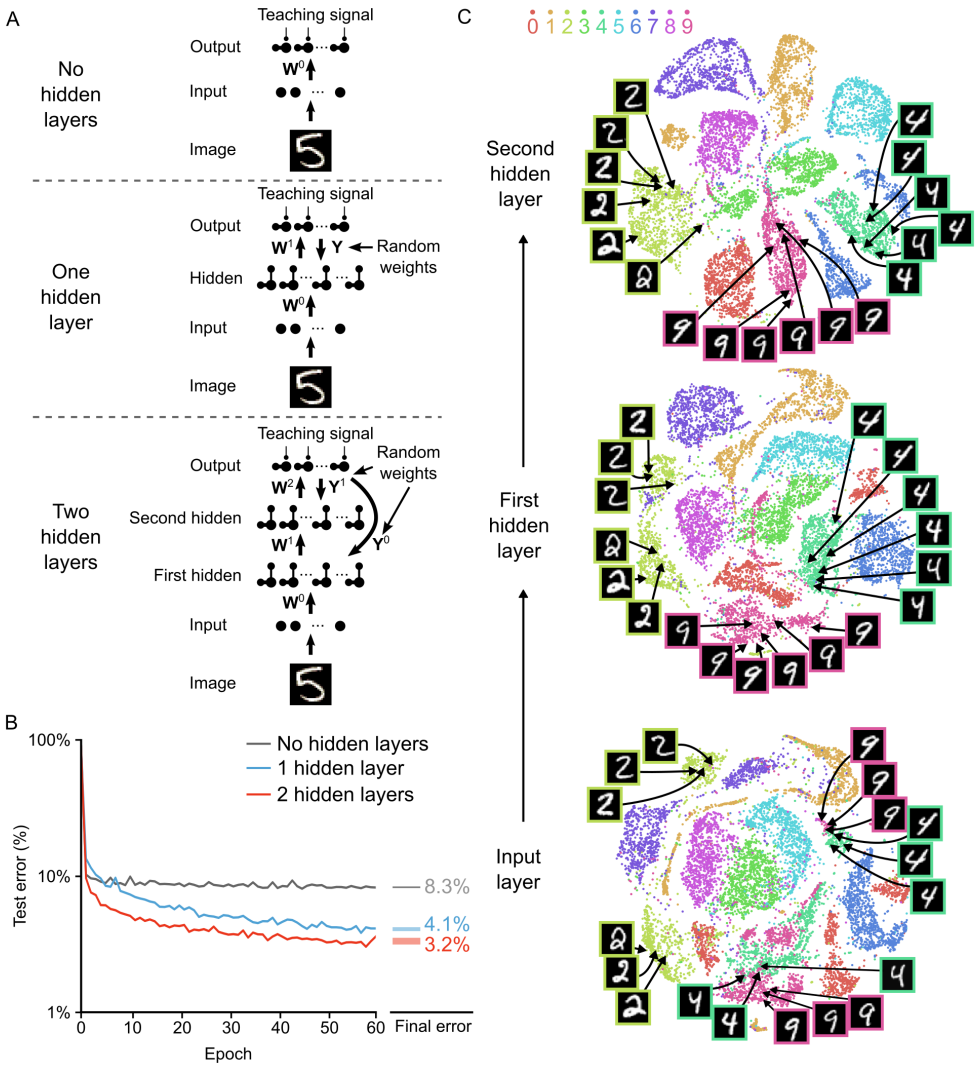

Figure 5. Improvement of learning with hidden layers.

(**A**) Illustration of the three networks used in the simulations. *Top*: a shallow network with only an input layer and an output layer. *Middle*: a network with one hidden layer. *Bottom*: a network with two hidden layers. Both hidden layers receive feedback from the output layer, but through separate synaptic connections with random weights $\boldsymbol{Y}^0$ and $\boldsymbol{Y}^1$. (**B**) Plot of test error (measured on 10,000 MNIST images not used for training) across 60 epochs of training, for all three networks described in **A**. The networks with hidden layers exhibit deep learning, because hidden layers decrease the test error. *Right*: Spreads (min – max) of the results of repeated weight tests ($n = 20$) after 60 epochs for each of the networks. Percentages indicate means (two-tailed t-test, 1-layer vs. 2-layer: $t_{38} = 197.11$, $P_{38} = 2.5 \times 10^{-58}$; 1-layer vs. 3-layer: $t_{38} = 238.26$, $P_{38} = 1.9 \times 10^{-61}$; 2-layer vs. 3-layer: $t_{38} = 42.99$, $P_{38} = 2.3 \times 10^{-33}$, Bonferroni correction for multiple comparisons). (**C**) Results of t-SNE dimensionality reduction applied to the activity patterns of the first three layers of a two hidden layer network (after 60 epochs of training). Each data point corresponds to a test image shown to the network. Points are color-coded according to the digit they represent. Moving up through the network, images from identical categories are clustered closer together and separated from images of different catefories. Thus the hidden layers learn increasingly abstract representations of digit categories.

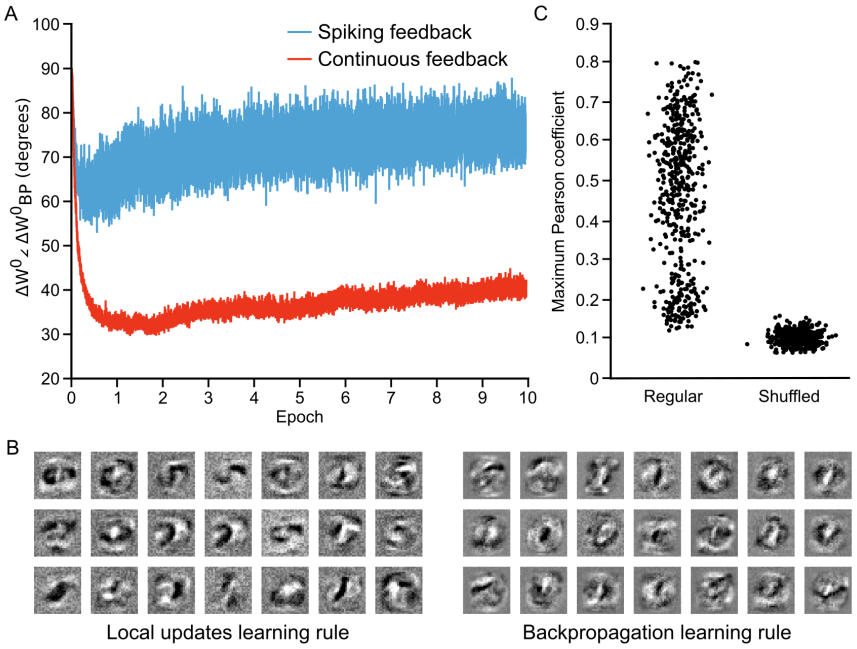

Figure 6. Approximation of backpropagation with local learning rules.

(**A**) Plot of the angle between weight updates prescribed by our local update learning algorithm compared to those prescribed by backpropagation of error, for a one hidden layer network over 10 epochs of training (each point on the horizontal axis corresponds to one image presentation). Data was time-averaged using a sliding window of 100 image presentations. When training the network using the local update learning algorithm, feedback was sent to the hidden layer either using spiking activity from the output layer units (blue) or by directly sending the spike rates of output units (red). The angle between the local update $\Delta \boldsymbol{W}^0$ and backpropagation weight updates $\Delta \boldsymbol{W}^0_{BP}$ remains under 90° during training, indicating that both algorithms point weight updates in a similar direction. (**B**) Examples of hidden layer receptive fields (synaptic weights) obtained by training the network in **A** using our local update learning rule (left) and backpropagation of error (right) for 60 epochs. (**C**) Plot of correlation between local update receptive fields and backpropagation receptive fields. For each of the receptive fields produced by local update, we plot the maximum Pearson correlation coefficient between it and all 500 receptive fields learned using backpropagation (Regular). Overall, the maximum correlation coefficients are greater than those obtained after shuffling all of the values of the local update receptive fields (Shuffled).

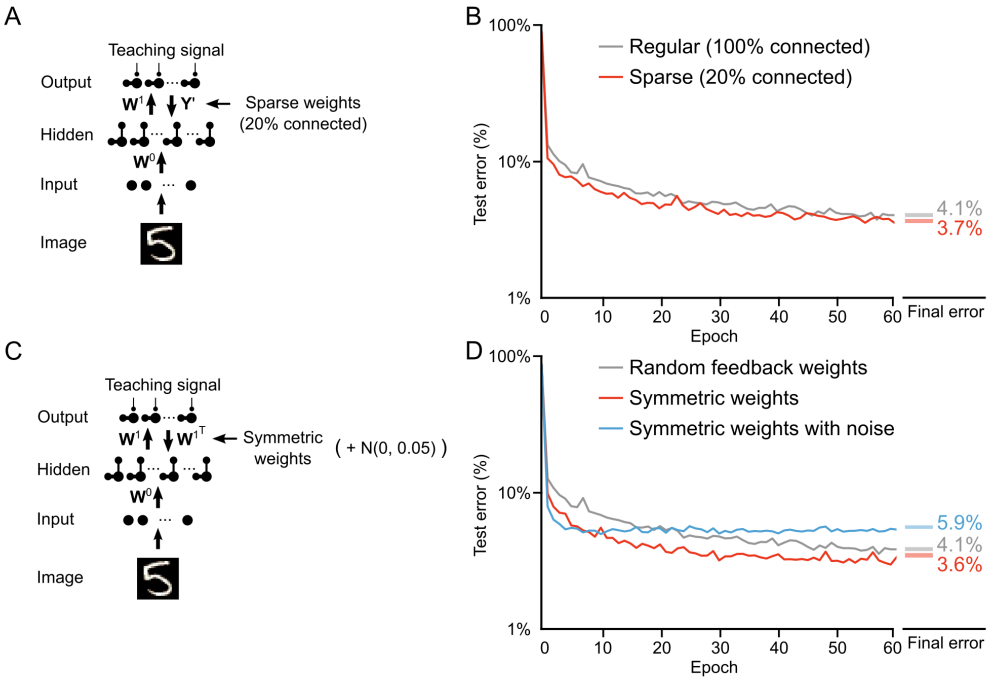

Figure 7. Conditions on feedback synapses for effective learning.

(**A**) Diagram of a one hidden layer network trained in **B**, with 80% of feedback weights set to zero. The remaining feedback weights $\boldsymbol{Y}'$ were multiplied by 5 in order to maintain a similar overall magnitude of feedback signals. (**B**) Plot of test error across 60 epochs for our standard one hidden layer network (gray) and a network with sparse feedback weights (red). Sparse feedback weights resulted in improved learning performance compared to fully connected feedback weights. *Right*: Spreads (min – max) of the results of repeated weight tests ($n = 20$) after 60 epochs for each of the networks. Percentages indicate mean final test errors for each network (two-tailed t-test, regular vs. sparse: $t_{38} = 16.43$, $P_{38} = 7.4 \times 10^{-19}$). (**C**) Diagram of a one hidden layer network trained in **D**, with feedback weights that are symmetric to feedforward weights $\boldsymbol{W}^1$, and symmetric but with added noise. Noise added to feedback weights is drawn from a normal distribution with variance $\sigma = 0.05$. (**D**) Plot of test error across 60 epochs of our standard one hidden layer network (gray), a network with symmetric weights (red), and a network with symmetric weights with added noise (blue). Symmetric weights result in improved learning performance compared to random feedback weights, but adding noise to symmetric weights results in impaired learning. *Right*: Spreads (min – max) of the results of repeated weight tests ($n = 20$) after 60 epochs for each of the networks. Percentages indicate means (two-tailed t-test, random vs. symmetric: $t_{38} = 18.46$, $P_{38} = 4.3 \times 10^{-20}$; random vs. symmetric with noise: $t_{38} = -71.54$, $P_{38} = 1.2 \times 10^{-41}$; symmetric vs. symmetric with noise: $t_{38} = -80.35$, $P_{38} = 1.5 \times 10^{-43}$, Bonferroni correction for multiple comparisons).

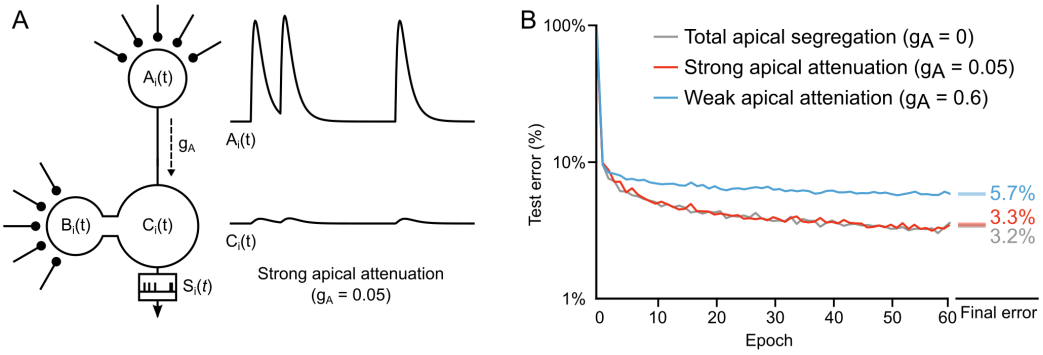

Figure 8. Importance of dendritic segregation for deep learning.

(**A**) *Left*: Diagram of a hidden layer neuron with an unsegregated apical dendrite that affects the somatic potential C_i. g_A represents the strength of the coupling between the apical dendrite and soma. *Right*: Example traces of the apical voltage A_i and the somatic voltage C_i in response to spikes arriving at apical synapses. Here $g_A = 0.05$, so the apical activity is strongly attenuated at the soma. (**B**) Plot of test error across 60 epochs of training on MNIST of a two hidden layer network, with total apical segregation (gray), strong apical attenuation (red) and weak apical attenuation (blue). Apical input to the soma did not prevent learning if it was strongly attenuated, but weak apical attenuation impaired deep learning. *Right*: Spreads (min – max) of the results of repeated weight tests ($n = 20$) after 60 epochs for each of the networks. Percentages indicate means (two-tailed t-test, total segregation vs. strong attenuation: $t_{38} = -4.00$, $P_{38} = 8.4 \times 10^{-4}$; total segregation vs. weak attenuation: $t_{38} = -95.24$, $P_{38} = 2.4 \times 10^{-46}$; strong attenuation vs. weak attenuation: $t_{38} = -92.51$, $P_{38} = 7.1 \times 10^{-46}$, Bonferroni correction for multiple comparisons).

Figure Supplements

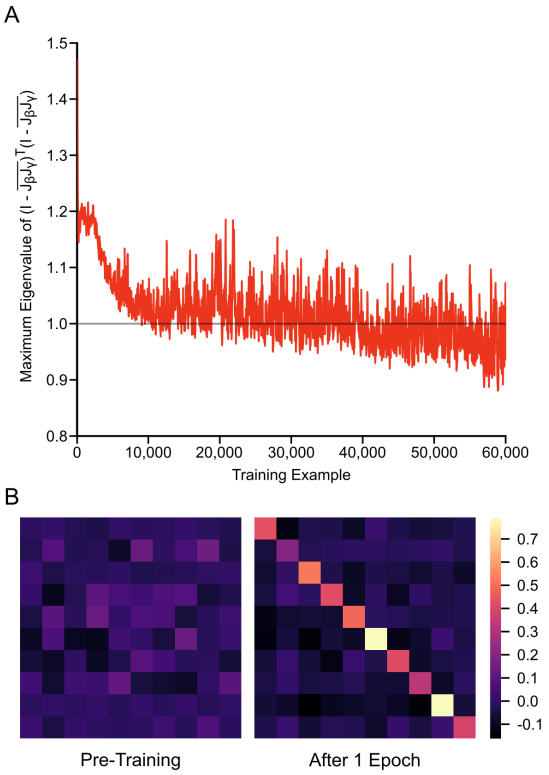

Figure 4, Supplement 1. Weight alignment during first epoch of training.

(**A**) Plot of the maximum eigenvalue of $(I - \overline{J_\beta J_\gamma})^T (I - \overline{J_\beta J_\gamma})$ over 60,000 training examples for a one hidden layer network, where $\overline{J_\beta}$ and $\overline{J_\gamma}$ are the mean feedforward and feedback Jacobian matrices for the last 100 training examples. The maximum eigenvalue of $(I - \overline{J_\beta J_\gamma})^T (I - \overline{J_\beta J_\gamma})$ drops below 1 as learning progresses, satisfying the main condition for the learning guarantee described in Theorem 1 to hold. (**B**) The product of the mean feedforward and feedback Jacobian matrices, $\overline{J_\beta J_\gamma}$, for a one hidden layer network, before training (left) and after 1 epoch of training (right). As training progresses, the network updates its weights in a way that causes this product to approach the identity matrix, meaning that the two matrices are roughly inverses of each other.

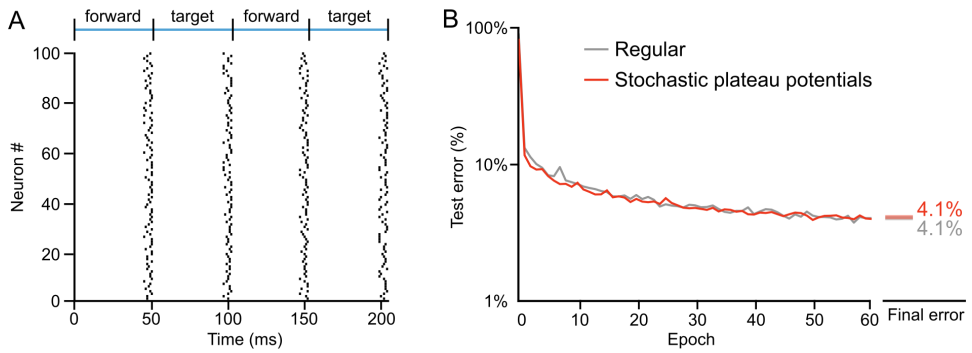

Figure 5, Supplement 1. Learning with stochastic plateau times.

(**A**) *Left*: Raster plot showing plateau potential times during presentation of two training examples for 100 neurons in the hidden layer of a network where plateau potential times were randomly sampled for each neuron from a folded normal distribution ($\mu = 0, \sigma^2 = 3$) that was truncated (max = 5) such that plateau potentials occurred between 0 ms and 5 ms before the start of the next phase. In this scenario, the apical potential over the last 30 ms was integrated to calculate the plateau potential for each neuron. (**B**) Plot of test error across 60 epochs of training on MNIST of a one hidden layer network, with synchronized plateau potentials (gray) and with stochastic plateau potentials (red). Allowing neurons to undergo plateau potentials in a stochastic manner did not hinder training performance.

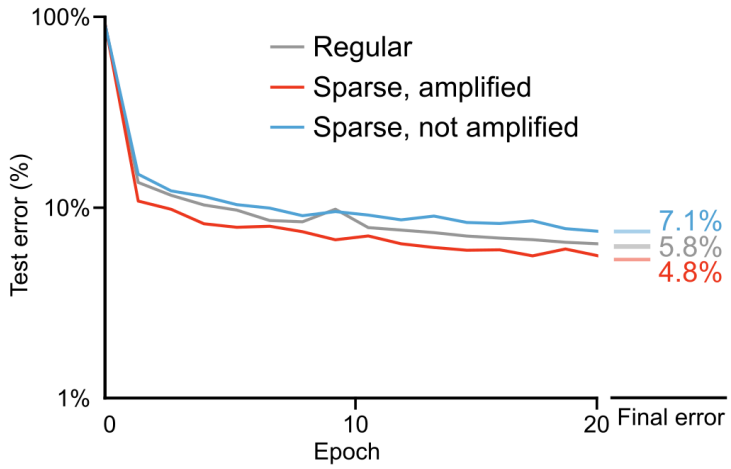

Figure 7, Supplement 1. Importance of weight magnitudes for learning with sparse weights.

Plot of test error across 20 epochs of training on MNIST of a one hidden layer network, with regular feedback weights (gray), sparse feedback weights that were amplified (red), and sparse feedback weights that were not amplified (blue). The network with amplified sparse feedback weights is the same as in Figure 7A & B, where feedback weights were multiplied by a factor of 5. While sparse feedback weights that were amplified led to improved training performance, sparse weights without amplification impaired the network's learning ability. *Right*: Spreads (min – max) of the results of repeated weight tests ($n = 20$) after 20 epochs for each of the networks. Percentages indicate means (two-tailed t-test, regular vs. sparse, amplified: $t_{38} = 44.96$, $P_{38} = 4.4 \times 10^{-34}$; regular vs. sparse, not amplified: $t_{38} = -51.30$, $P_{38} = 3.2 \times 10^{-36}$; sparse, amplified vs. sparse, not amplified: $t_{38} = -100.73$, $P_{38} = 2.8 \times 10^{-47}$, Bonferroni correction for multiple comparisons).