Neuron
Report

Learning by the Dendritic Prediction of Somatic Spiking

Robert Urbanczik[1] and Walter Senn[1,*]
[1]Department of Physiology, and Center for Learning, Cognition and Memory, University of Bern, Bühlplatz 5, CH-3012 Bern, Switzerland
*Correspondence: senn@pyl.unibe.ch
http://dx.doi.org/10.1016/j.neuron.2013.11.030

SUMMARY

Recent modeling of spike-timing-dependent plasticity indicates that plasticity involves as a third factor a local dendritic potential, besides pre- and postsynaptic firing times. We present a simple compartmental neuron model together with a non-Hebbian, biologically plausible learning rule for dendritic synapses where plasticity is modulated by these three factors. In functional terms, the rule seeks to minimize discrepancies between somatic firings and a local dendritic potential. Such prediction errors can arise in our model from stochastic fluctuations as well as from synaptic input, which directly targets the soma. Depending on the nature of this direct input, our plasticity rule subserves supervised or unsupervised learning. When a reward signal modulates the learning rate, reinforcement learning results. Hence a single plasticity rule supports diverse learning paradigms.

1. INTRODUCTION

In spike-timing-dependent plasticity (STDP) experiments, potentiation is only observed when pre- and postsynaptic spike pairs are induced with a sufficiently high frequency (Markram et al., 1997). This, by itself, has long since indicated that neurons that fire together do not unconditionally wire together. But a comprehensive phenomenological model of such non-Hebbian effects has only recently been achieved by including voltage as a third modulating factor for plasticity, in addition to the traditional pre-/posttimings (Clopath and Gerstner, 2010; Clopath et al., 2010). In vivo, the modulating voltage is thought to correspond to a local dendritic potential. But it can be estimated by low-pass filtering the somatic potential, in the special case that action potentials are elicited by somatic current injection, as in the classical STDP experiments.

Theoretical studies on the function of STDP have mostly assumed point neurons (Abbott and Nelson, 2000; Song et al., 2000; Kempter et al., 2001; Gütig et al., 2003). But this seems inadequate, if plasticity is modulated by a local dendritic voltage, which in vivo may substantially differ from the somatic potential. Here we present a compartmental neuron model and derive from first principles a plasticity rule in which the voltage modulation of synaptic plasticity has a functional interpretation. Remarkably, plasticity becomes simpler for this more complex, but arguably more realistic, model neuron, in that a single learning rule now encompasses diverse learning paradigms.

In designing the compartmental model, our overarching goal was simplicity, since we want to retain the key advantages of point neuron models: amenability to analytical insight and usability in large-scale simulations. As a consequence, our model rides roughshod over many aspects of neuronal morphology and dynamics. For instance, we collapse the complex neuronal morphology into a single somatic and a single dendritic compartment. Further, subthreshold voltage in our model propagates from the dendrite to the soma but not vice versa. However, simulation results indicate that the important functional aspects of our model and plasticity rule do not depend crucially on these simplifying assumptions (Supplemental Information available online).

A key aspect of voltage dependence is that in the subthreshold regime the strength of the synaptic depression resulting from an unpaired presynaptic input increases with voltage (Artola et al., 1990; Clopath and Gerstner, 2010). This leads us to conceptualize plasticity in dendritic synapses as driving a predictive coding scheme that adapts the dendritic potential to match the somatic activity. The likelihood of a somatic spike should increase with increasing dendritic input; hence, when there is no somatic spike in spite of high dendritic voltage, the synapses that caused the elevated voltage get strongly depressed. Conversely, a somatic spike that is unexpected due to a relatively low value of the dendritic potential gives rise to potentiation. Hence, we propose that plasticity is not driven by the correlation between pre- and postsynaptic activity, as in Hebbian learning, but by the correlation of presynaptic activity with a postsynaptic, somatodendritic prediction error.

Deviations between dendritic potential and somatic activity, i.e., prediction errors, can result in our model from stochastic fluctuations as well as from synaptic input that directly targets the soma. Depending on the source of the direct input into the soma, the proposed plasticity rule implements supervised or unsupervised learning. The stochastic fluctuations lead to exploratory somatic activity that can subserve reinforcement learning. Indeed, when there is no direct synaptic input to the soma, our plasticity rule becomes mathematically equivalent to a rule previously derived in a point neuron model (Pfister et al., 2006), which has since been widely used in reinforcement learning (Di Castro et al., 2009; Urbanczik and Senn, 2009; Frémaux et al., 2010). As a consequence, the proposed model provides a unified plasticity rule for unsupervised, supervised, and reinforcement learning.

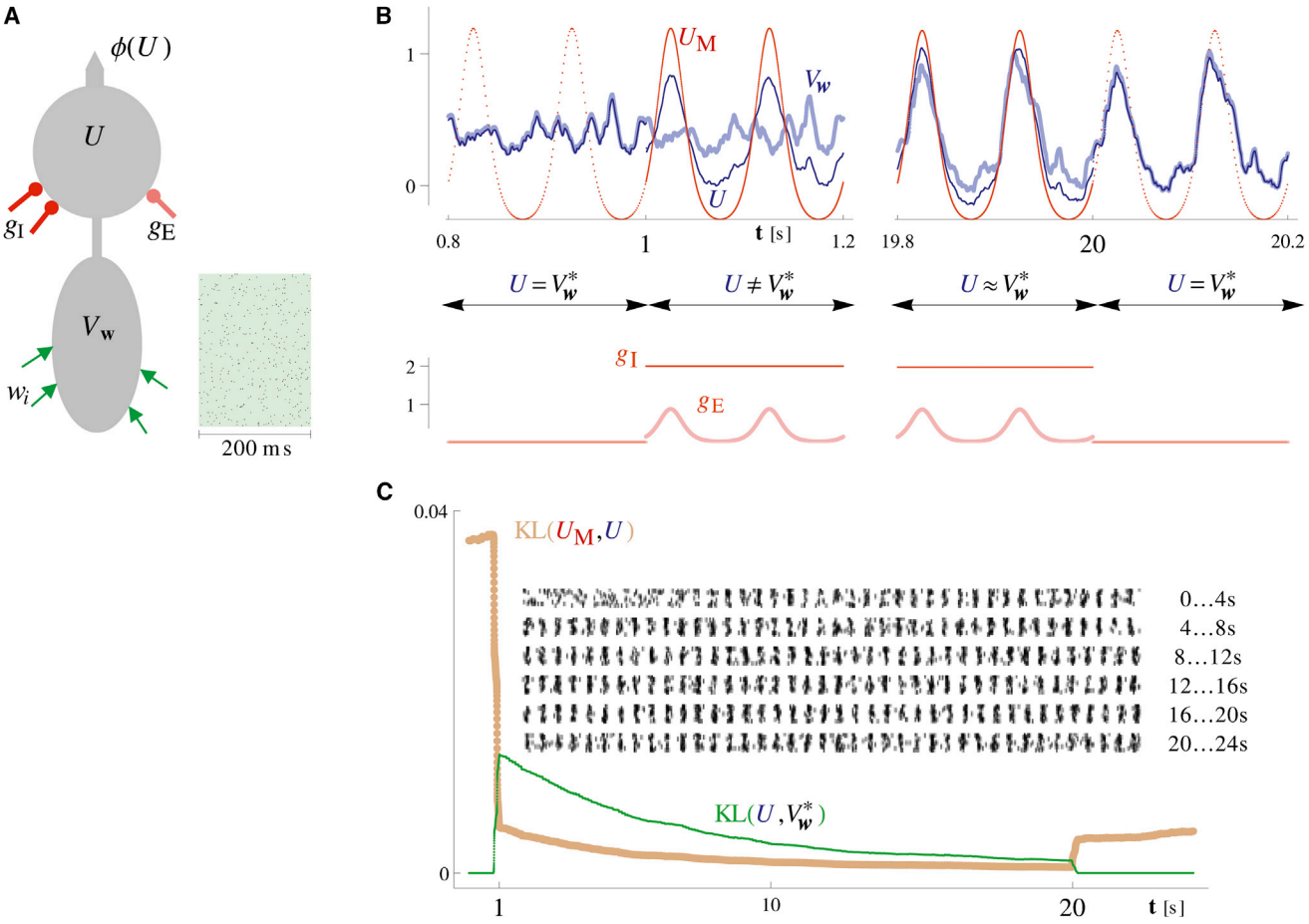

Figure 1. Learning in the Compartmental Neuron

(A) Sketch of the neuronal model and a typical dendritic input pattern. (B) Traces of key neuronal variables. During the entire run, the input spike pattern shown in (A) is presented over and over again. Starting at $t = 1$ s, the nudging conductances g_E and g_I are active for 19 s, encoding U_M (filled red curve, top row) as the target time course for the somatic potential. This target time course is shown by the red dotted curve, at times when the nudging conductances are off. (C) Average learning curves for the above task, based on $n = 10$ runs with a different input pattern and different initial dendritic weights for each run. $KL(U, V_w^*)$ assesses the discrepancy between actual somatic firings and firings predicted by the dendritic potential. It is calculated by using a statistical measure (Kullback-Leibler divergence, Supplemental Information, Equation S12) to compare the firing rates $\phi(U)$ and $\phi(V_w^*)$. The plot for $KL(U_M, U)$) shows the discrepancy between target firings (rate $\phi(U_M)$) and actual firings. In the nudging phase, dendritic plasticity by decreasing $KL(U, V_w)$ also reduces $KL(U_M, U)$. Hence $KL(U_M, U)$ stays small when the nudging is inactivated at $t = 20$ s. Inset: the somatic spikes produced during the ten runs.

2. RESULTS

We first describe our model and the plasticity rule in the context of a simple supervised learning task, a mathematical derivation of the rule as gradient procedure is given in Supplemental Information, Section 2.1. In this first task, a single neuron learns to associate a dendritic input with a somatic target response. For learning, somatic activity is modulated by the target response via somatic synapses and this leads to prediction errors. After learning, the target response is produced solely from the dendritic input even when the somatic synapses are silent. A next simulation demonstrates how this learning principle subserves the formation of associative memories in a recurrent network: during learning, each memory pattern is delivered to the network by the somatic synapses, and each neuron learns to produce its component of the pattern from the dendritic synapses connecting the neuron to the other neurons in the network. Unsupervised learning arises in our model when the somatic synapses are driven by the learning network itself. For this, a last simulation shows how our plasticity rule can lead to the self-organization of a topographic mapping (Kohonen, 1982).

2.1. The Model in a Simple Supervised Learning Task

The model neuron that we consider here (Figure 1A) is made up of a somatic and a single dendritic compartment, but it can be extended to more than one dendritic compartment (Supplemental Information, Figure S1).

Somatic Compartment

Depending on value of the somatic potential U, the soma generates spikes probabilistically as in a Poisson process but with a 3 ms refractory period. In particular, its instantaneous firing rate is a sigmoidal function $\phi(U)$ of the somatic voltage. The

somatic compartment integrates input from the dendritic compartment with inputs from proximal synapses (current I_U^{som}). Hence, the potential U evolves as

$$\dot{U} = -g_L U + g_D(V_{\mathbf{w}} - U) + I_U^{\text{som}} \qquad \text{(Equation 1)}$$

where we have set the capacitance C to unity and omitted this factor. Further, g_L is the leak conductance, g_D the coupling of the dendrite to the soma, and $V_{\mathbf{w}}$ the dendritic voltage. The subscript $\mathbf{w}$ in $V_{\mathbf{w}}$ refers to the vector of the synaptic strengths in the dendrite. For the input from the synapses proximal to the soma, we adopt the conductance-based formulation

$$I_U^{\text{som}}(t) = g_E(t)(E_E - U) + g_I(t)(E_I - U). \qquad \text{(Equation 2)}$$

Here g_E and g_I are time-varying excitatory and inhibitory conductances with reversal potentials E_E and E_I.

Importantly, the balance of excitation and inhibition defines an effective reversal potential for which proximal synaptic input does not generate any current. We call this reversal potential the matching potential U_M. By setting Equation 2 to zero it is obtained as

$$U_M(t) = \frac{g_E(t)E_E + g_I(t)E_I}{g_E(t) + g_I(t)}. \qquad \text{(Equation 3)}$$

Note that U_M does indeed act as a reversal potential: a value of U smaller than U_M results in $I_U^{\text{som}}>0$ but, conversely, the direct synaptic input current is negative if the somatic voltage lies above the matching potential. So the proximal synapses nudge the somatic potential toward U_M (Figure 1B for t between 1 and 1.2 s), and we will refer to g_E and g_I as the nudging conductances.

Dendritic Prediction

Let us for a moment assume that the proximal nudging synapses are silent ($g_E = g_I = 0$), as in Figure 1B for $t<1$ s. Then the time course of the somatic voltage is determined just by the dendrite and shall, in this case, be denoted by $V_{\mathbf{w}}^*$. One can easily calculate $V_{\mathbf{w}}^*$ by integrating Equation 1 under the condition $I_U^{\text{som}} = 0$. Here we shall assume a strong coupling g_D of the soma to the dendrite and then $V_{\mathbf{w}}^* = (gD/(gD + gL))V_{\mathbf{w}}$ holds to an excellent approximation. So in the absence of proximal somatic input, the somatic potential is in essence a slightly attenuated version of the dendritic potential (Figure 1B for t between 0.8 and 1 s). Despite the simple relationship to $V_{\mathbf{w}}$, the value of $V_{\mathbf{w}}^*$ is important conceptually because the soma would fire with rate $\phi(V_{\mathbf{w}}^*)$ if somatic synapses were always silent. We interpret the notional rate $\phi(V_{\mathbf{w}}^*)$ as the dendritic prediction of the actual somatic firing rate $\phi(U)$ and conceptualize learning in the dendritic synapses as aiming to reduce the rate prediction error, that is the magnitude of $\phi(U) - \phi(V_{\mathbf{w}}^*)$.

Rate prediction errors can arise from direct synaptic input to the soma that nudges the somatic potential U away from $V_{\mathbf{w}}^*$. Crucially, nudging must not always lead to such errors, since no current flow arises from the nudging when the somatic potential equals the matching potential U_M. So if the dendritic potential by itself follows a time course such that $V_{\mathbf{w}}^* = U_M$, the nudging has no effect. In this sense, the dendrite can predict away the proximal synaptic input. Approximately, this is the case in Figure 1B for t between 19.8 and 20 s.

When rate prediction errors do arise from the nudging, dendritic plasticity reducing these errors sets into motion a virtuous cycle in which the somatic potential U serves as an intermediate moving target. An adaptive change to $V_{\mathbf{w}}$ reducing the magnitude of $\phi(U) - \phi(V_{\mathbf{w}}^*)$ moves $V_{\mathbf{w}}$ toward U. Since U lies in between $V_{\mathbf{w}}^*$ and the matching potential, the change also moves $V_{\mathbf{w}}^*$ closer to U_M. However, the adaptive change in the dendrite also influences the somatic potential and as a consequence U moves toward U_M. This change to U recreates a prediction error, triggering further adaptive change until the dendrite in the end catches up with the soma when both $V_{\mathbf{w}}^*$ and U converge to U_M. Hence, while the intermediate target of the learning process is U, the effective, final target is the matching potential U_M.

Plasticity Rule

For minimizing it, dendritic synapses need to estimate the rate prediction error $\phi(U) - \phi(V_{\mathbf{w}}^*)$. The predicted rate $\phi(V_{\mathbf{w}}^*)$ is readily obtained from the local dendritic potential $V_{\mathbf{w}}$. The actual somatic rate can be estimated based on the back propagation of action potentials, since the somatic spike train $S(t)$ provides a noisy observation process for the underlying firing rate $\phi(U(t))$. Statistically, $\phi(U(t))$ is the expectation of $S(t)$, when the spike train $S(t)$ is given as a sum of δ-functions centered at somatic spike times. So a noisy estimate of the rate prediction error is provided by $S(t) - \phi(V_{\mathbf{w}}^*(t))$, and we assume that it is this estimate that drives synaptic plasticity. For a dendritic synapse i with strength w_i, we introduce the plasticity induction variable PI_i by

$$\text{PI}_i(t) = \left(S(t) - \phi(V_{\mathbf{w}}^*(t))\right) h\left(V_{\mathbf{w}}^*(t)\right) \frac{\partial}{\partial w_i} V_{\mathbf{w}}(t). \qquad \text{(Equation 4)}$$

Here h is a positive weighting function, the choice of which we discuss below. The exact form of the partial derivative term $(\partial/\partial w_i)V_{\mathbf{w}}$ depends on the model for the dendritic compartment. Here, we adopt a simple spike response model, in which the dendritic voltage is given as a weighted sum of the spike response functions for the afferents (Equation 8, Experimental Procedures). The weight of each afferent is the synaptic strength w_i and its spike response functions $\text{PSP}_i(t)$ is determined solely by the presynaptic spike timings. Then the derivative is just $(\partial/\partial w_i)V_{\mathbf{w}} = \text{PSP}_i(t)$. So, as in the phenomenological model (Clopath and Gerstner, 2010; Clopath et al., 2010), plasticity induction is determined by three factors: pre-/posttiming and the dendritic potential.

Note that one can regard the difference $S(t) - \phi(V_{\mathbf{w}}^*(t))$ in Equation 4 as the instantaneous prediction error. Since $S(t)$ records the actual spiking of the neuron, this error is never zero; the neuron cannot produce, say, half a spike. But averaged over many trials, $S(t) - \phi(V_{\mathbf{w}}^*(t)$ converges to the rate prediction error $\phi(U) - \phi(V_{\mathbf{w}}^*)$, which can be zero. So even if plasticity is induced in every trial, the changes can cancel and then only negligible net synaptic change results. However, such trial-by-trial fluctuations can potentially be reinforced by a reward signal, and this provides the hook for using the rule in reinforcement learning. It is this scenario, which motivated our choice of the weight function h in Equation 4, namely $h(x) = (d/dx)\ln\phi(x)$. For silent nudging conductances, the above choice makes our plasticity model mathematically equivalent to a previously derived reinforcement learning rule (Pfister et al., 2006). For

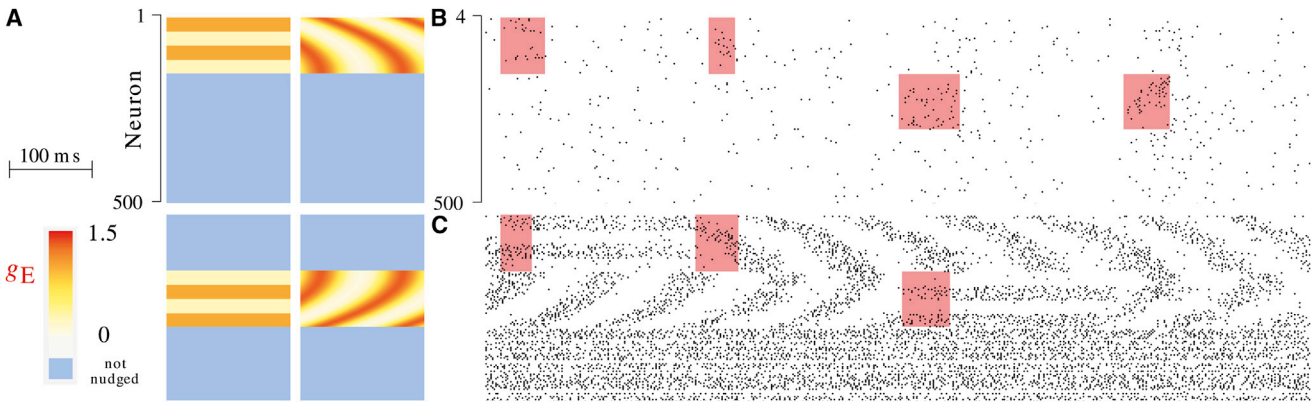

Figure 2. Memory Formation in a Network with Recurrent Somatodendritic Connections

The 100 ms scale bar refers to all panels. (A) Excitatory nudging conductances for the four patterns. When a neuron is nudged, the excitation is balanced by an inhibitory conductance of $g_I = 3$. For no nudging, $g_E = g_I = 0$. (B and C) Response to brief nudging: before learning (B) and after 500 s of learning (simulated biological time) (C). Neurons and times for which nudging occurs are marked by the red shading. For visibility, only every fourth neuron is shown. A statistical evaluation of recall performance is given in Supplemental Information, Section 3.3.

just supervised learning, there would be considerable leeway in choosing h, and one might even just omit this term (Supplemental Information, Equations S4 and S7).

In the model, induced plasticity is low-pass filtered with a time constant τ_Δ before being consolidated into persistent synaptic change

$$\tau_\Delta \dot{\Delta}_i = \mathrm{PI}_i(t) - \Delta_i$$
$$\dot{w}_i = \eta \Delta_i. \qquad \text{(Equation 5)}$$

Here η is the learning rate. For one thing, the low-pass filtering dampens the fluctuations arising from the noisy estimation of the rate prediction error. More importantly, it provides a time lag, making it possible to use the rule with a delayed reward signal in reinforcement learning by assuming that the learning rate η is proportional to an external reward signal. In this case, the mathematical equivalence of our model to the previous work on reinforcement learning is strict (Pfister et al., 2006; Frémaux et al., 2010); hence, we will only consider supervised and unsupervised learning in the rest of the paper.

Returning to supervised learning, Figure 1C shows the learning curves for the simple scenario considered in Figure 1B. The curves highlight that the plasticity rule moves the somatic potential toward the matching potential by reducing the somatodendritic prediction error resulting from the nudging. We emphasize that dendritic synapses are oblivious of whether or not the soma is being nudged and get updated throughout the entire session (Equations 4 and 5). But, whenever the nudging conductances are silent, the updates are random and the net change in the neuronal response stays negligible.

For nudging, inhibition and excitation play an opposing but symmetric role in our model. Although, in terms of conductance injected into the soma, excitatory nudging is weaker than inhibition, simply because excitatory conductance leads to strong current flow due to its high reversal potential. The broadly symmetric role is nevertheless at variance with findings on the prevalence of somatic inhibition in principal cells (Somogyi et al., 1998). In Supplemental Information, we show how the model can be adapted to take into account a high baseline level of somatic inhibition (Figure S2).

2.2. Associative Memory

As a more involved learning task, we consider memory formation in a network of 500 compartmental neurons. Recurrent connections relay the spikes of each neuron to the dendritic compartment of other neurons (50% random connectivity), and dendritic synapses follow the above plasticity rule. The four patterns that we trained with are shown in Figure 2A; two of the patterns use a rate code, the other two use a phase code. For the phase code patterns, pattern neurons have firing rate profiles with identical period but shifted phases. Similar patterns arise in the phase coding of path information observed in hippocampal place cells of rodents (O'Keefe and Recce, 1993; Huxter et al., 2003).

Patterns were imprinted on the network by randomly selecting one of the four patterns during learning and then nudging the pattern neurons for an average duration of 500 ms. Thereafter, the procedure was repeated with a next pattern. As a recall paradigm for the patterns, we consider the brief nudging of the pattern neurons. Before learning, the network responds to the recall nudgings by weak and brief activity (Figure 2B). After learning, the patterns are represented by sustained activity states and the nudging triggers transitions between the states (Figure 2C).

In Figure 2, some neurons are not part of any pattern and are thus never nudged. Similarly to echo state networks (Jaeger and Haas, 2004), such "hidden" neurons may enhance the representational capabilities of the network. Here, however, dendritic synapses are plastic in all neurons. So there is no biophysical difference between "hidden" and "visible" neurons and a neuron's role in the network is assigned ad hoc during learning depending on whether it happens to get nudged.

2.3. Self-Organized Topographic Mappings

Till now we have not considered the source of the direct somatic input. Taking this into account becomes crucial, however,

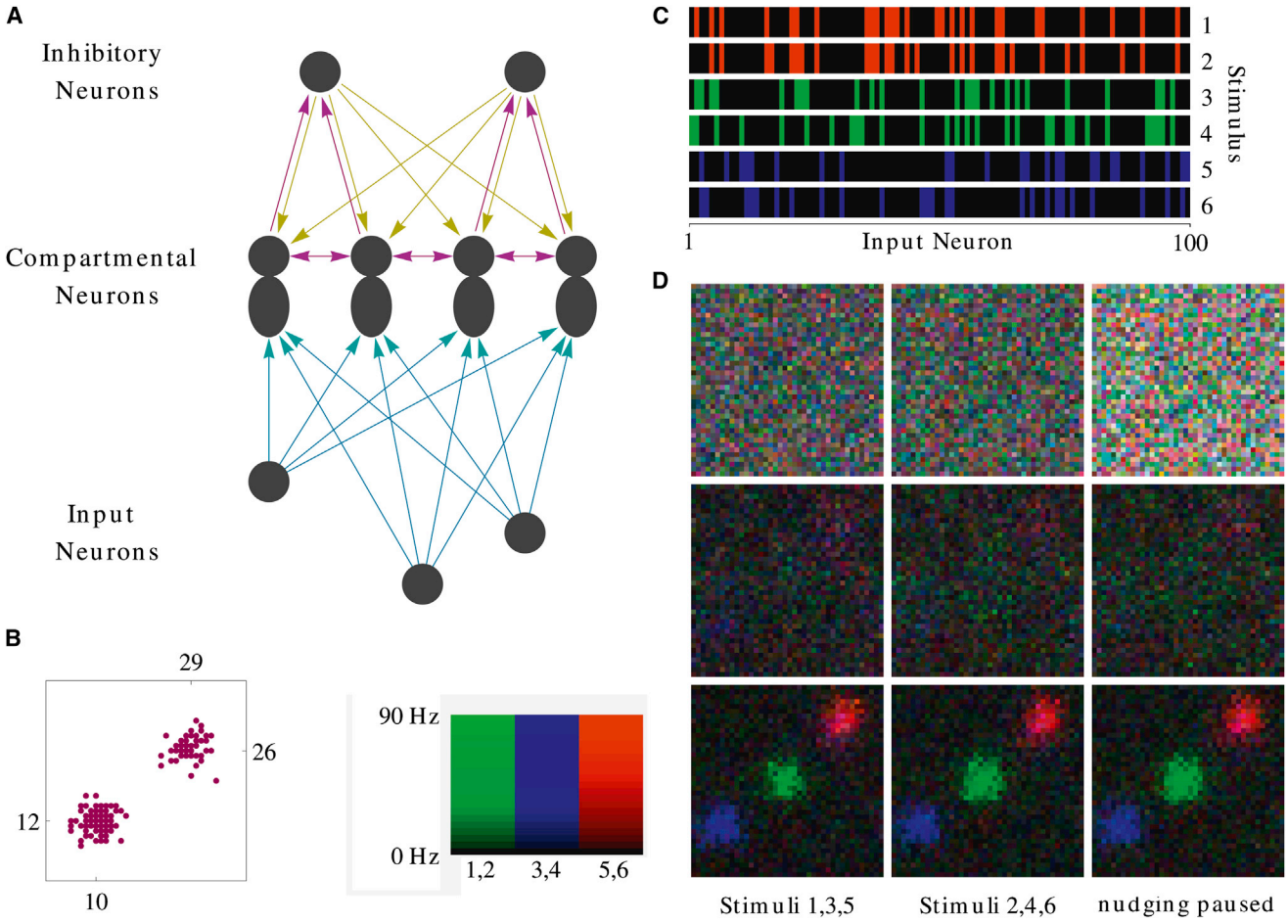

Figure 3. Learning a Topographic Mapping

(A) Sketch of a one-dimensional network analogous to the two-dimensional topographic network that we used. The color coding of the connections is: cyan, plastic; magenta, excitatory; yellow, inhibitory. (B) Sample of the excitatory somatosomatic connectivity in the network that we actually used. Dots mark the neurons with lateral input from neuron (12, 10) and neuron (29, 26). (C) Poisson firing rates of the input neurons to the network, the six stimuli form three color-coded clusters. (D) Network response before learning (top row), after 700 s of simulated biological learning time (middle row), and after 1,600 s of learning (bottom row). Mean firing rates recorded during a 1 s presentation of each stimulus are shown. The left column records the responses to the first input stimulus in each of the clusters; the middle column is for the second stimuli. For the right column, stimuli 2, 4, 6 were presented as in the middle column. However, just during the 1 s activity recording, all somatosomatic interaction was turned off to show the stimulus response in the absence of nudging. Firing rates in (B) and (C) are color coded as indicated by the legend. In (C), responses to different stimuli are overlayed by additively combining their color codes. More simulation details are given in Supplemental Information, Section 3.4.

whenever the network itself generates the nudging. Then the supervised learning in effect turns into an unsupervised learning procedure since the network computes its own teaching signal. As an example, we modeled the learning of a two-dimensional topographic mapping by our plasticity rule. A sketch of a one-dimensional network, analogous to the one we used, is shown in Figure 3A. Weak short-range somatosomatic connections provide excitatory nudging and this is balanced by a weak long-range inhibitory nudging that is mediated by a pool of inhibitory interneurons. To promote stimulus selectivity, the interneurons have facilitating afferents (see Zucker and Regehr, 2002; Supplemental Information, Section 3.4). The network that we simulated had 40×40 compartmental and 20×20 inhibitory neurons. A sample of the actual excitatory somatosomatic connectivity is shown in Figure 3B.

The dendrites of the compartmental neurons are connected to an input layer presenting on each trial one of the stimuli shown in Figure 3C. The six rate-coded stimuli have no manifest topographical organization, but the stimulus set does have some structure. It is made up of three clusters (color coded), with strong within but weak between cluster correlations. Mapping these stimuli topographically yields activity patterns that are similar to the spatially clustered object representations observed in inferotemporal cortex (Tanaka, 2003).

Recordings of stimuli responses in the compartmental neurons are shown in Figure 3D. Activity is initially disorganized but learning leads to a spatial organization in which the different clusters in the input are mapped to different areas in the network. Importantly, topographic organization emerges from learning and not from the recurrent network dynamics. This is highlighted

by the right column in Figure 3D, in which all nudging was turned off during activity recording in order to obtain the feedforward response of the network. Comparing to the middle column shows that the stimulus response is hardly influenced by the recurrences except for a small, globally inhibitory influence before learning.

A similar network using point neurons and strong recurrent connections was presented for this task by Michler et al. (2009). There even before learning, due to the strong recurrences, network activity shows substantial topographic organization that is then further enhanced by assuming an appropriate plasticity rule for the feedforward connections: potentiation uses a Hebbian mechanism and this is balanced by depression implemented as a multiplicative synaptic scaling, modulated by the postsynaptic firing frequency. In contrast, when plasticity is driven by a prediction error, there is no need to assume that plasticity is tailored to the unsupervised learning task. In our case, the balance of synaptic potentiation and depression results from, and reflects, the balance of excitation and inhibition in the somatosomatic connectivity. In essence, the unsupervised learning algorithm is not specified by the plasticity rule but by the architecture of the network.

3. DISCUSSION

Our model of self-organized feature maps (Figure 3) shows that, by shaping plasticity, even weak synaptic input can have a profound long-term effect on processing. This provides an angle on longstanding discussions on the relative importance of fast feedforward versus slower recurrent processing in the brain (Lamme and Roelfsema, 2000; Ganguli and Latham, 2009). In vision, for instance, Riesenhuber and Poggio (2000) have argued that the observed fast reaction times (Stanford et al., 2010) are suggestive of a processing that is predominantly feedforward. This, however, seems at odds with the massive recurrent connectivity in visual cortex. Our simulation with weak but persistent nudging highlights the possibility that some of the recurrence could subserve the learning of appropriate feedforward mappings, even if the recurrent connections only marginally affect the instantaneous stimulus response.

Supervised learning depends on the distinction between a target value and an actually produced value. Modeling this with point neurons is awkward because of the difficulty of fitting two values into a single point. Since the seminal work of Hopfield (1982) on associative memory, one has typically assumed that time multiplexing provides a solution and distinguished a learning phase from a retrieval phase. During learning, the neuronal output is "clamped" to the target without being affected by the adapting afferents for which Hebbian plasticity is turned on. During retrieval, the neuronal output is driven by the adapted afferents, but plasticity is now turned off to keep the neuron from learning any mistakes it might make. Recently a more subtle version of time multiplexing has been suggested, in which the target value is delivered to the neuron with a precisely timed delay after the actual value has been produced (D'Souza et al., 2010).

With two compartments, supervised learning is much simpler. The nudging of the somatic compartment provides information on the target value, whereas the dendrite produces the actual value. When learning is driven by the somatodendritic rate prediction error, net plasticity induction decreases by learning and becomes zero as soon as the nudging stops. Further, after arriving at the target defined by the matching potential, nudging can be turned on or off without affecting the somatic potential since the conductance-based somatic input becomes ineffective once its postsynaptic reversal potential is reproduced by appropriately learned dendritic input. In effect, successful learning explains away the teacher and there is thus no need for a temporally precise control of plasticity that distinguishes between learning and retrieval phases. A key requirement for this, at the level of the dendritic synapses, is the modulation of plasticity by dendritic voltage. Currently, evidence for such a modulation is circumstantial (Artola et al., 1990; Clopath and Gerstner, 2010). It would be desirable to have more proximal data, obtained by patching the dendrite in the vicinity of the synapse being investigated in order to control the local voltage.

While average plasticity induction is zero in the absence of rate prediction errors, plasticity is nevertheless induced all of the time in our model, driven by instantaneous prediction errors that have zero average over many trials. So when there is no need for learning, there is no plasticity induction on average, and it seems hard to imagine a better way of reconciling the stability of learned associations with ongoing synaptic plasticity. But even if this may be the best solution to the stability-plasticity dilemma, it is not a perfect solution. A close inspection of Figure 1C shows a small deterioration in performance during the 4s of learning after $t = 20$ when the nudging stopped. This arises because in the absence of nudging, ongoing plasticity causes the synaptic strength to evolve as in a random walk. In Supplemental Information (Figure S5), we show that such synaptic diffusion can go on for 10 to 20 s without causing dramatic changes to the learned neuronal behavior. Further, in a network of densely connected neurons, changes on the single neuron level will tend to cancel. Also, no plasticity at all is induced by our rule when there is no activity. While these two mechanisms can lead to learned associations persisting quite a bit longer, it is hard to imagine them resulting in persistence on the order of days. Hence, even in our model, some explicit control of plasticity is needed to guarantee the stability of learned associations. This chimes in with the experimental findings that so-called long-term potentiation or depression is not tantamount to lasting potentiation or depression (Pastalkova et al., 2006; Frey and Frey, 2008). Instead, newly induced changes to synaptic strength are labile for a time frame of hours and whether they become persistent can even depend on seemingly unrelated behavioral events (Ballarini et al., 2009). Such memory consolidation mechanisms are suggestive of the temporally coarse-grained control of plasticity needed for the persistence of learned associations in our model.

EXPERIMENTAL PROCEDURES

Full simulation details are given in the Supplemental Information (Section 3). Here the complete description of the model neuron is presented, starting with the dendritic compartment.

In the dendrite, we adopt for simplicity a synaptic model that is not conductance based. Instead, presynaptic input directly and immediately leads to

current injection into the dendritic compartment. So presynaptic spike trains, X_i^{dnd}, jointly give rise to a dendritic input current I^{dnd} evolving as

$$\tau_s \dot{I}^{\text{dnd}} = -I^{\text{dnd}} + \sum_i w_i \sum_{s \in X_i^{\text{dnd}}} \delta(t-s) \qquad \text{(Equation 6)}$$

where we think of each spike train X_i^{dnd} as the set of the presynaptic spike times in afferent i. The synaptic strength for this afferent is w_i, and we use $\tau_s = 3$ for the synaptic time constant. Here, and throughout the Experimental Procedures, we measure time in milliseconds.

The dendritic potential $V_{\mathbf{w}}$ is obtained by low-pass filtering the input current using:

$$\tau_L \dot{V}_{\mathbf{w}} = -V_{\mathbf{w}} + I^{\text{dnd}} \qquad \text{(Equation 7)}$$

with $\tau_L = 10$ for the leak time constant. Equation 7 can be solved analytically, resulting in the spike response form of $V_{\mathbf{w}}$. Then the dendritic voltage is obtained as

$$V_{\mathbf{w}}(t) = \sum_i w_i \text{PSP}_i(t) \text{ where } \text{PSP}_i(t) = \sum_{s \in X_i^{\text{dnd}}} \kappa(t-s) \qquad \text{(Equation 8)}$$

with the response kernel $\kappa(t) = (1/(\tau_L - \tau_s))\Theta(t)(e^{-t/\tau_L} - e^{-t/\tau_s})$. We emphasize that this is a highly stylized model of a dendritic compartment. In particular, Equation 7 does not allow for any current flow from soma to dendrite. In the Supplemental Information, we show that this restriction can be relaxed (Figures S3 and S4)

The presynaptic term in our learning rule (Equation 4) is the partial derivative $(\partial/\partial w_i)V_{\mathbf{w}}(t)$. From (Equation 8) this results in $(\partial/\partial w_i)V_{\mathbf{w}}(t) = \text{PSP}_i(t)$, the sum of the response kernel κ over presynaptic spike times. It is this simple form that made us choose the spike response model for the dendrite, instead of using a more realistic conductance-based formulation. While calculating the partial derivative with respect to a synaptic strength in a conductance-based model is straightforward, the derivative depends on the total amount of conductance in the dendrite because a high level of conductance, in effect, increases the leak. We are not aware of any experimental results regarding such a nonlocal modulation of the presynaptic term. In a conductance-based model, one can probably get away with using an approximate gradient based on a standard response kernel for the purpose of simulations. But this by itself would make it difficult to make contact with any mathematical theory.

As mentioned in the Results (Equation 1), the somatic potential evolves as

$$\dot{U}(t) = -g_L U(t) + g_D(V_{\mathbf{w}}(t) - U(t)) + I_U^{\text{som}}(t)$$

where we use $g_L = 1/\tau_L = 0.1$ for the leak conductance and $g_D = 2$ for the coupling of the dendrite to the soma. Note that for both $V_{\mathbf{w}}$ and U the resting potential is 0. The somatic current is given by $I_U^{\text{som}}(t) = g_E(t)(E_E - U(t)) + g_I(t)(E_I - U(t))$, with $E_E = 4.667$ and $E_I = -1/3$. Below, we will choose a soft spiking threshold of $\vartheta = 1$. If one takes our unitless resting potential of 0 to correspond to -70 mV, and our unitless threshold of 1 to correspond to -55 mV, the above choices for E_E and E_I correspond to reversal potentials of 0 mV (excitation) and -75 mV (inhibition).

When explicitly modeling how the nudging conductances g_E and g_I arise from the firing of other neurons synapsing onto the soma, we assume that a presynaptic spike leads to an instantaneous increase in conductance followed by an exponential decay. For the total excitatory and inhibitory conductance in the soma, this results in

$$\begin{aligned} \dot{g}_E &= -g_E/\tau_s + \sum_j w_j^E \sum_{s \in X_j^E} \delta(t-s) \\ \dot{g}_I &= -g_I/\tau_s + \sum_k w_k^I \sum_{s \in X_k^I} \delta(t-s). \end{aligned} \qquad \text{(Equation 9)}$$

Here w_j^E is the strength of the j-th excitatory synapse proximal to the soma and X_j^E its presynaptic input spike train. For inhibition, the homologous role is played by w_k^I and X_k^I.

In the mathematical analysis (Supplemental Information, Section 2.1), we assume Poisson spiking for the soma with an instantaneous rate $\phi(U(t))$, i.e., the probability of generating a spike in the time interval $[t, t+\delta t]$ is $\phi(U(t))\delta t$ in the limit of small δt. For biological realism, we modified this slightly in the simulations presented here, by assuming a 3 ms absolute refractory period after each spike, during which the soma cannot generate further spikes. For the rate function ϕ, we use a sigmoidal of the form

$$\phi(U) = \frac{\phi_{\max}}{1 + k e^{\beta(\vartheta - U)}}. \qquad \text{(Equation 10)}$$

with $\phi_{\max} = 0.15$, $k = 0.5$, $\beta = 5$ and, as mentioned above, $\vartheta = 1$. The choice of $\phi_{\max}$ means that the maximal firing rate is 0.15 kHz. Note that this maximal rate is attained only for an infinite value of U, maximal firing rates in our simulations are considerably lower. Assuming a sigmoidal for $\phi(U)$ is not an essential part of our model, in principle any rate function that increases with U could be used (Supplemental Information, Section 2.1).

Taking the definition of $V_{\mathbf{w}}$ into account (Equation 8), the equation for plasticity induction (Equation 4) becomes

$$\text{PI}_i(t) = \Big(S(t) - \phi\big(V_{\mathbf{w}}^*(t)\big)\Big) h\Big(V_{\mathbf{w}}^*(t)\Big) \text{PSP}_i(t) \qquad \text{(Equation 11)}$$

with $h(x) = (\text{d}/\text{d}x) \ln \phi(x)$. Equation 11 depends on the dendritic potential but we do not explicitly model the back propagation of somatic action potentials into the dendrite. Hence our theory is incapable of describing any voltage dependence of plasticity induction while an action potential is ongoing. We take this into account by not using Equation 11 during the refractory period immediately after a spike. In this period, the soma will not spike whatever the dendrite does, so we simply set $\text{PI}_i(t) = 0$ during this period. In practical terms, given the typical firing rates in our simulations, we expect the effects of refractoriness to be minor. For instance, in Figure 1 the mean value of our performance measure $\text{KL}(U_M, U)$ immediately after learning is 0.0037 ± 0.0003. This value changes to 0.0031 ± 0.0003 when the simulation is rerun with no refractoriness.

As mentioned in the main text, $\text{PI}(t)$ is low-pass filtered before inducing synaptic change using: $\tau_\Delta \dot{\Delta}_i = \text{PI}_i(t) - \Delta_i$ and $\dot{w}_i = \eta \Delta_i$. In the simulations, $\tau_\Delta = 100$. The learning rate η is different for the different tasks (Supplemental Information, Section 3).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and five figures and can be found with this article online at http://dx.doi.org/10.1016/j.neuron.2013.11.030.

ACKNOWLEDGMENTS

This research is supported by the Swiss National Science Foundation (SNF, Grant 31003A_133094) and a grant from the Swiss SystemsX.ch initiative (Neurochoice), evaluated by the SNF.

Accepted: November 21, 2013
Published: February 5, 2014

REFERENCES

Abbott, L.F., and Nelson, S.B. (2000). Synaptic plasticity: taming the beast. Nat. Neurosci. Suppl. *3*, 1178–1183.

Artola, A., Bröcher, S., and Singer, W. (1990). Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex. Nature *347*, 69–72.

Ballarini, F., Moncada, D., Martinez, M.C., Alen, N., and Viola, H. (2009). Behavioral tagging is a general mechanism of long-term memory formation. Proc. Natl. Acad. Sci. USA *106*, 14599–14604.

Clopath, C., and Gerstner, W. (2010). Voltage and spike timing Interact in STDP - a unified model. Fron.t Synaptic Neurosci. *2*, 25.

Clopath, C., Büsing, L., Vasilaki, E., and Gerstner, W. (2010). Connectivity reflects coding: a model of voltage-based STDP with homeostasis. Nat. Neurosci. *13*, 344–352.

D'Souza, P., Liu, S.C., and Hahnloser, R.H. (2010). Perceptron learning rule derived from spike-frequency adaptation and spike-time-dependent plasticity. Proc. Natl. Acad. Sci. USA *107*, 4722–4727.

Di Castro, D., Volkinshtein, S., and Meir, R. (2009). Temporal difference based actor critic learning - convergence and neural implementation. In Advances in Neural Information Processing Systems, *Volume 21*, D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, eds. (Cambridge: MIT Press), pp. 385–392.

Frémaux, N., Sprekeler, H., and Gerstner, W. (2010). Functional requirements for reward-modulated spike-timing-dependent plasticity. J. Neurosci. *30*, 13326–13337.

Frey, S., and Frey, J.U. (2008). 'Synaptic tagging' and 'cross-tagging' and related associative reinforcement processes of functional plasticity as the cellular basis for memory formation. Prog. Brain Res. *169*, 117–143.

Ganguli, S., and Latham, P. (2009). Feedforward to the past: the relation between neuronal connectivity, amplification, and short-term memory. Neuron *61*, 499–501.

Gütig, R., Aharonov, R., Rotter, S., and Sompolinsky, H. (2003). Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. J. Neurosci. *23*, 3697–3714.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. USA *79*, 2554–2558.

Huxter, J., Burgess, N., and O'Keefe, J. (2003). Independent rate and temporal coding in hippocampal pyramidal cells. Nature *425*, 828–832.

Jaeger, H., and Haas, H. (2004). Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. Science *304*, 78–80.

Kempter, R., Gerstner, W., and van Hemmen, J.L. (2001). Intrinsic stabilization of output rates by spike-based Hebbian learning. Neural Comput. *13*, 2709–2741.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. Biol. Cybern. *43*, 59–69.

Lamme, V.A., and Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci. *23*, 571–579.

Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science *275*, 213–215.

Michler, F., Eckhorn, R., and Wachtler, T. (2009). Using spatiotemporal correlations to learn topographic maps for invariant object recognition. J. Neurophysiol. *102*, 953–964.

O'Keefe, J., and Recce, M.L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. Hippocampus *3*, 317–330.

Pastalkova, E., Serrano, P., Pinkhasova, D., Wallace, E., Fenton, A.A., and Sacktor, T.C. (2006). Storage of spatial information by the maintenance mechanism of LTP. Science *313*, 1141–1144.

Pfister, J.P., Toyoizumi, T., Barber, D., and Gerstner, W. (2006). Optimal spike-timing-dependent plasticity for precise action potential firing in supervised learning. Neural Comput. *18*, 1318–1348.

Riesenhuber, M., and Poggio, T. (2000). Models of object recognition. Nat. Neurosci. Suppl. *3*, 1199–1204.

Somogyi, P., Tamás, G., Lujan, R., and Buhl, E.H. (1998). Salient features of synaptic organisation in the cerebral cortex. Brain Res. Brain Res. Rev. *26*, 113–135.

Song, S., Miller, K.D., and Abbott, L.F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nat. Neurosci. *3*, 919–926.

Stanford, T.R., Shankar, S., Massoglia, D.P., Costello, M.G., and Salinas, E. (2010). Perceptual decision making in less than 30 milliseconds. Nat. Neurosci. *13*, 379–385.

Tanaka, K. (2003). Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. Cereb. Cortex *13*, 90–99.

Urbanczik, R., and Senn, W. (2009). Reinforcement learning in populations of spiking neurons. Nat. Neurosci. *12*, 250–252.

Zucker, R.S., and Regehr, W.G. (2002). Short-term synaptic plasticity. Annu. Rev. Physiol. *64*, 355–405.

Supplemental Information

Learning by the dendritic prediction of somatic spiking

Robert Urbanczik, Walter Senn

Department of Physiology, University of Bern, Bühlplatz 5, CH-3012 Bern, Switzerland

---

Contents

1 Variations on the model

While arguably an improvement over point neurons, our compartmental model still gives a highly stylized account of biological reality. One issue is that few neuronal morphologies are well described in terms of a somatic compartment and a single dendritic compartment. Hence in the next subsection we show how the model can be adapted when there is more than one dendritic compartment.

A second issue arises from the fact that synaptic input directly targeting the soma has often been found to be largely inhibitory. In a limited sense, the simulations in the main text are in line with this because in terms of conductance

strength the inhibitory nudging is always considerably stronger than the excitatory nudging. This reflects the fact that the gap between reversal potential and spiking threshold is much bigger for excitation than for inhibition. Hence for physiological values of the somatic potential, excitatory conductance is more powerful than inhibitory conductance.

Our model can however be adapted to the case where there is an even larger discrepancy between somatic excitation and inhibition. This is shown in Section 1.2 where the dendritic prediction is modified to take into account a baseline of somatic inhibition.

A further issue is our assumption that subthreshold current flows from the dendrite to the soma, but not in the reverse direction. Partial justification for this is provided by passive cable theory, since the theory posits that it is the ratio of the surface areas of the two compartments which determines the ratio between the coupling constant g_{D} for current flow from dendrite to soma and the coupling constant g_{S} for the reverse flow. For many neuronal morphologies the dendritic surface is much larger than that of the soma, suggesting that $g_{\mathrm{D}}/g_{\mathrm{S}}$ can be large. But large is not infinite, as was the case in main text, and hence we present simulation results for our model with a nonzero value of g_{S}. Finally we show how the plasticity rule can even be adapted to the case that g_{D} is not large.

1.1 Two dendritic compartments

We assume two dendritic compartments with local potentials $V_{\mathbf{w}}^{(1)}$ and $V_{\mathbf{w}}^{(2)}$. Each compartment integrates its presynaptic input just like the single compartment in the main text. To aggregate the two dendritic compartments in the soma, we replace Eq. (1) in the main text by

$$\dot{U} = -g_{\mathrm{L}}U + g_{\mathrm{D}}(V_{\mathbf{w}}^{(1)} - U) + g_{\mathrm{D}}(V_{\mathbf{w}}^{(2)} - U) + I_U^{\mathrm{som}}$$

Each compartment now makes its own prediction $V_{\mathbf{w}}^{(i)*}$ for the somatic potential and we set $V_{\mathbf{w}}^{(i)*} = \frac{2g_{\mathrm{D}}}{2g_{\mathrm{D}}+g_{\mathrm{L}}}V_{\mathbf{w}}^{(i)}$. The plasticity rule for a synapse in dendritic compartment i now uses $V_{\mathbf{w}}^{(i)*}$ as the modulation factor instead of $V_{\mathbf{w}}^{*}$ (Eq. 4 and 11, main text). Our choice for the dendritic predictions means that in addition to assuming that there is no nudging, each compartment assumes that its prediction of the somatic potential is the same as the prediction of the other compartment. While this assumption is incorrect before learning, we expect the voltage discrepancies between the compartments to decrease through learning since their overall influence on the soma is determined by the mean (or, more generally, a convex combination) of their local dendritic potentials.

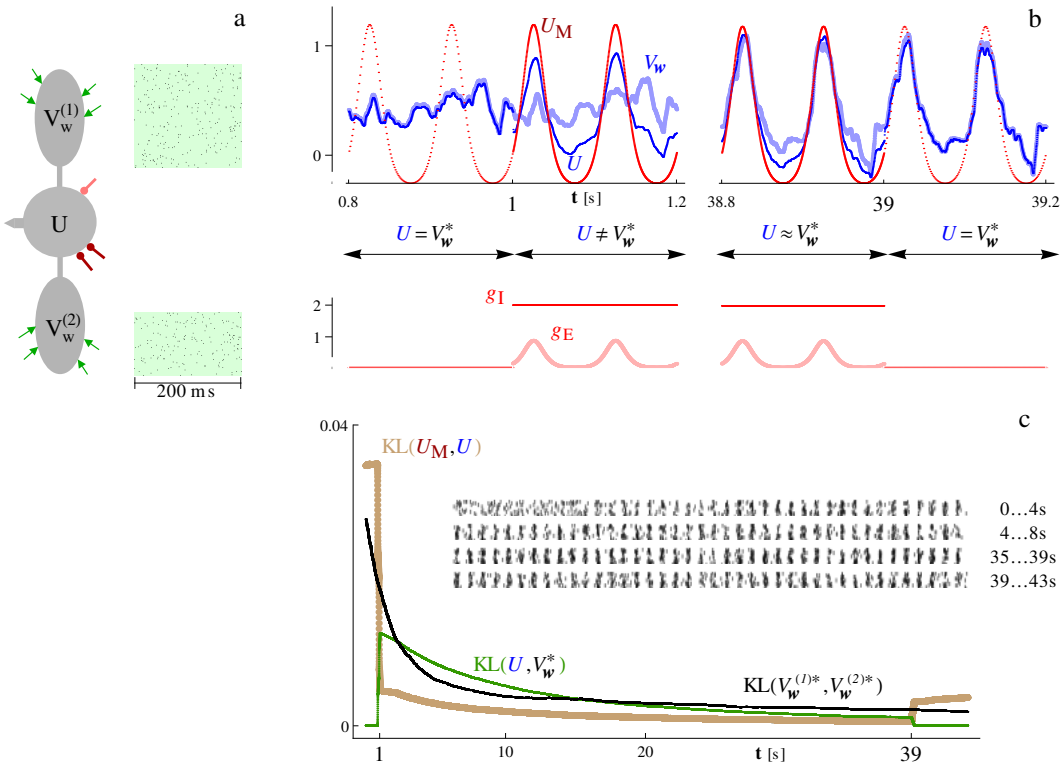

Fig. S1. Learning with two dendritic compartments. The learning scenario and the information shown is entirely analogous to Fig. 1 in the main text. In (**b**) we introduce as a purely descriptive quantity the mean $V_{\mathbf{w}}$ of $V_{\mathbf{w}}^{(1)}$ and $V_{\mathbf{w}}^{(2)}$, and we use $V_{\mathbf{w}}^* = \frac{2g_D}{2g_D+g_L} V_{\mathbf{w}}$ in (**c**).

Simulation results for the same task as in Fig. 1 of the main text are shown in Fig. S1. For the wiring, we assumed that the first 60% of the input neurons project to the first dendritic compartment and that the remaining input neurons project to the second dendritic compartment. Despite the asymmetry, the predictions of the two dendritic compartments converge during learning as shown by the $\mathrm{KL}(V_{\mathbf{w}}^{(1)*}, V_{\mathbf{w}}^{(2)*})$ curve in Panel S1c. Compared to the case of a single dendritic compartment, learning is slower since the dendritic predictions are less reliable. So we increased learning time by a factor of 2. To account for the fact that there a two dendritic compartments, the coupling constant g_D was decreased from its value of $g_D = 2$ in the main text to $g_D = 1$. Since each of the two dendritic compartments now has roughly half as many inputs, we doubled the initial values for the strength of the dendritic synapses. All other parameters are the same as in the corresponding simulation in the main text.

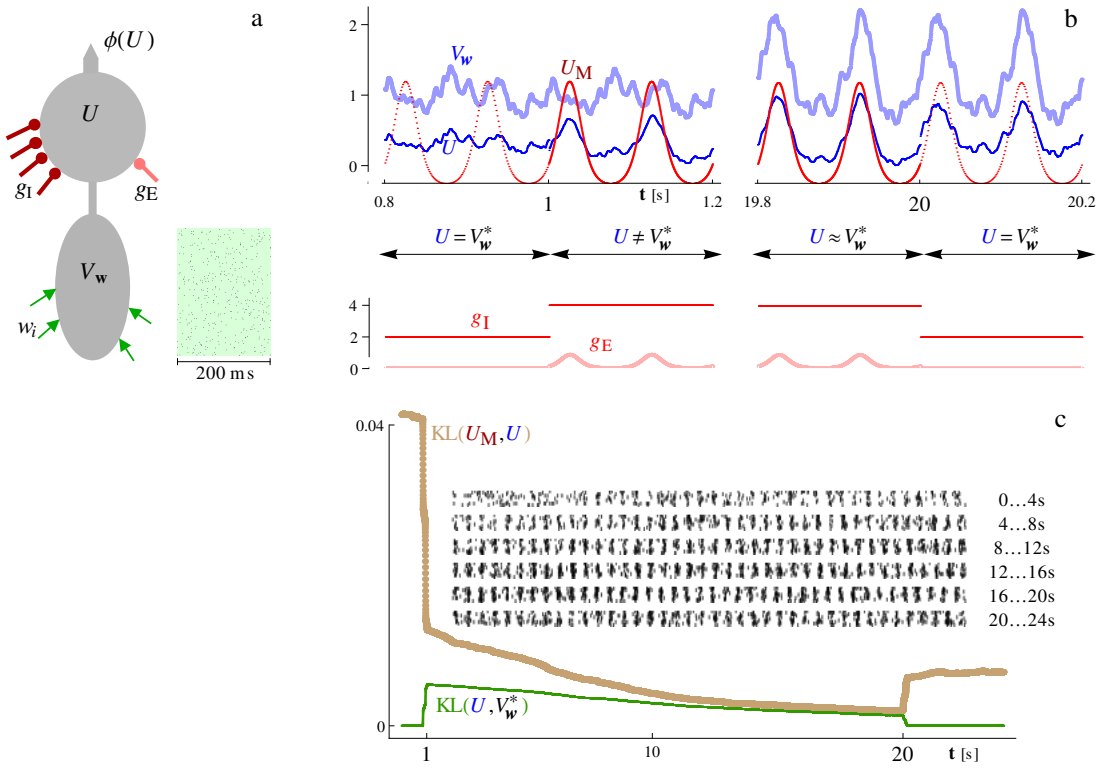

Fig. S2. Learning with a somatic baseline inhibition $g_I^{base} = 2$. The learning scenario and the information shown is entirely analogous to Fig. 1 in the main text. The calculation of the matching potential U_M is based on the expression in the main text (Eq. 3) with the following modification: Instead of using g_I for the inhibitory conductance, one has to use $g_I^{nudge} = g_I - g_I^{base}$.

1.2 Somatic baseline inhibition

Here we assume that the soma always receives inhibitory somatic input. To take this into account we consider the conductance g_I to be made up of two components g_I^{base} and g_I^{nudge}, so that $g_I = g_I^{base} + g_I^{nudge}$. A non-zero value of g_I^{base} corresponds to a constant baseline of inhibition in the soma which should not drive plasticity in the dendrite. Plasticity, however, should arise when the somatic voltage is nudged due to non-zero values of g_I^{nudge} and g_E.

From Eqs. (1,2) in the main text, we obtain that in the absence of nudging ($g_I^{nudge} = g_E = 0$) the somatic potential evolves as

$$\dot{U} = -g_L U + g_D(V_\mathbf{w} - U) + g_I^{base}(E_I - U)\,.$$

It is now this equation which has to be used for computing the dendritic prediction $V_\mathbf{w}^*$ of the somatic potential. We again assume that g_D is much larger than the leak conductance g_L, but not necessarily larger than the baseline

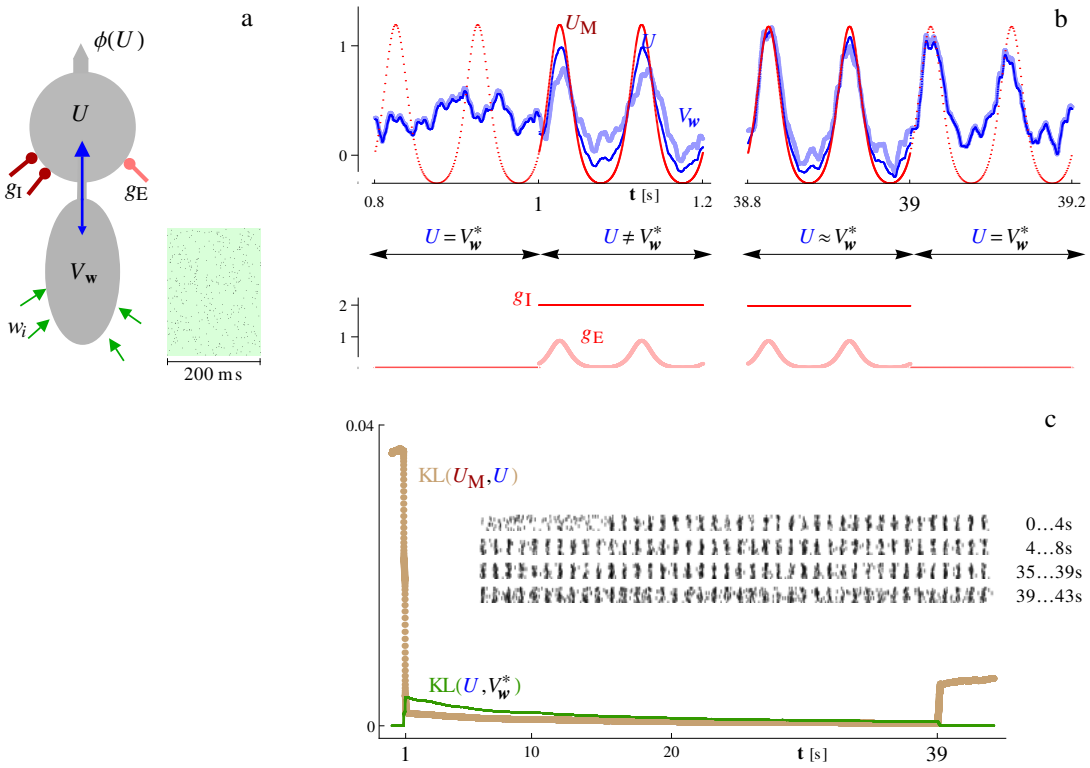

Fig. S3. Learning in the presence of subthreshold flow from soma to dendrite. The learning scenario and the information shown is entirely analogous to Fig. 1 in the main text.

inhibition g_I^{base}. Then to an excellent approximation

$$V_{\mathbf{w}}^* = \frac{g_D V_{\mathbf{w}} + g_I^{base} E_I}{g_D + g_I^{base} + g_L}$$

and it is this expression which has to be used for $V_{\mathbf{w}}^*$ in the plasticity rule (Eq. 4 and 11, main text) instead of the original $V_{\mathbf{w}}^* = \frac{g_D V_{\mathbf{w}}}{g_D + g_I}$. The rescaling of the dendritic prediction is the only change in the model in order to accommodate a prevalence of somatic inhibition.

The simulation results in Fig. S2 highlight the strong depolarization in the dendrite now needed to counteract the somatic inhibition. Learning performance, however, is similar to the case without baseline inhibition. Due to the increased inhibition, we chose larger initial synaptic strengths in the dendrite, picking them from a Gaussian distribution with a mean and a standard deviation of 1/2. We also increased the learning rate to $\eta = 0.1$. All other parameters are the same as in the simulation for Fig. 1 in the main text.

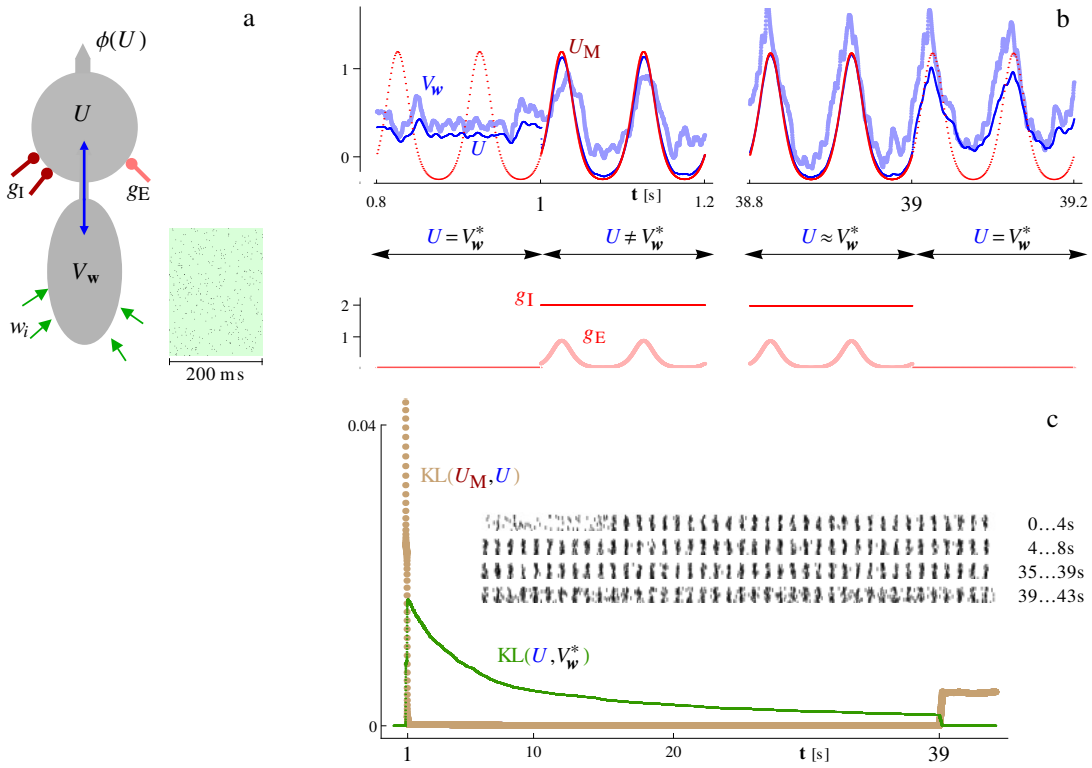

Fig. S4. Learning with a symmetrical coupling of soma and dendrite. The learning scenario and the information shown is entirely analogous to Fig. 1 in the main text.

1.3 Subthreshold flow from soma to dendrite

To account for the passive propagation from the soma to the dendrite we change Eq. (7) in Methods to

$$\dot{V}_{\mathbf{w}} = -V_{\mathbf{w}}/\tau_{\mathrm{L}} + g_{\mathrm{S}}(U - V_{\mathbf{w}}) + I^{\mathrm{dnd}}/\tau_{\mathrm{L}}$$

with $g_{\mathrm{S}} = 0.2$. As the plot in the left column of Fig. S3b shows, nudging the soma now has a significant effect on the dendritic potential. The remaining results in Fig. S3 show that this slows down but does not disrupt learning. Aside of the above change to the equation for $V_{\mathbf{w}}$ and the increase in learning time, all simulation details are the same as for Fig. 1 in the main text.

1.4 Symmetrical coupling of soma and dendrite

Here we assume $g_{\mathrm{D}} = g_{\mathrm{S}} = 0.2$. The small value of g_{D} means that propagation from the dendrite to the soma results not only in increased attenuation of $V_{\mathbf{w}}$ but also in a noticeable time lag. Hence in this section we no longer use the

attenuation approximation $V_{\mathbf{w}}^* \approx \frac{g_{\mathrm{D}} V_{\mathbf{w}}}{g_{\mathrm{D}}+g_{\mathrm{I}}}$ but obtain V_w^* by low pass filtering $V_{\mathbf{w}}$. So V_w^* is the solution of

$$\dot{U} = -g_{\mathrm{L}} U + g_{\mathrm{D}}(f - U) \tag{S1}$$

for $f(t) = V_{\mathbf{w}}(t)$. Similarly, we need to take the lag into account for the presynaptic term, so we let PSP_i^* be the solution of (S1) for $f(t) = \mathrm{PSP}_i(t)$. The modified plasticity rule then reads

$$\mathrm{PI}_i(t) = \Big(S(t) - \phi(V_{\mathbf{w}}^*(t))\Big)\, h\left(V_{\mathbf{w}}^*(t)\right) \mathrm{PSP}_i^*(t)\,.$$

The results with this rule in Fig. S4 show that learning is not disrupted by the symmetrical coupling, even if it is a bit slower than for Fig. 1 in the main text. The learning rate was $\eta = 0.15$ and a mean value of 0.3 was used for initializing the dendritic weights. All other simulation parameters where as in the preceding subsection.

2 Mathematical Analysis

In this section we return to the basic model considered in the main text, but assume for simplicity that there is no refractory period, i.e. somatic spiking is an inhomogeneous Poisson process with rate $\phi(U(t))$.

2.1 The plasticity rule as a gradient ascent procedure

We first combine the two equations defining the somatic potential (Eq. 1&2, main text) into the single equation

$$\dot{U} = -g_{\mathrm{tot}} U + g_{\mathrm{D}} V_{\mathbf{w}} + g_{\mathrm{E}} E_{\mathrm{E}} + g_{\mathrm{I}} E_{\mathrm{I}} \quad \text{with } g_{\mathrm{tot}} = g_{\mathrm{L}} + g_{\mathrm{D}} + g_{\mathrm{E}} + g_{\mathrm{I}}\,. \tag{S2}$$

We now consider the limit that the coupling g_{D} of the soma to dendrite is strong. In particular we assume that g_{D} is much larger than g_{L}, but make no assumption about the magnitude of g_{D} relative to the nudging conductances g_{E} and g_{I}. Expanding the solution of (S2) in powers of $1/g_{\mathrm{D}}$ we obtain

$$U = (1-\lambda) U_{\mathrm{M}} + \lambda V_{\mathbf{w}}^* + \mathcal{O}(\tfrac{1}{g_{\mathrm{D}}}) \quad \text{with } \lambda = \tfrac{g_{\mathrm{D}}+g_{\mathrm{L}}}{g_{\mathrm{tot}}}\,. \tag{S3}$$

So to leading order the somatic potential is a convex combination of the matching potential (U_{M}, Eq. 3 main text) and the dendritic prediction $V_{\mathbf{w}}^* = \frac{g_{\mathrm{D}}}{g_{\mathrm{D}}+g_{\mathrm{L}}} V_{\mathbf{w}}$. Note that the mixing factor λ will usually be time varying, when the nudging conductances and thus g_{tot} depend on time.

We now introduce the class of objective functions on which we want to do gradient ascent, namely

$$c_\lambda(u,v) = \int_0^v \mathrm{d}\nu \; h(\nu) \left(\phi\left((1-\lambda)u + \lambda\nu \right) - \phi(\nu) \right) . \tag{S4}$$

In the context of our model, λ is going to be the above mixing factor, ϕ the firing rate and h the weight function introduced in Eq. (4) of the main text. But for the moment just assume that $0 < \lambda < 1$, that ϕ is monotonically increasing and that h is positive. For the partial derivative with respect to v of the cost function (S4 we have

$$\tfrac{\partial}{\partial v} c_\lambda(u,v) = h(v) \left(\phi\left((1-\lambda)u + \lambda v \right) - \phi(v) \right) \tag{S5}$$

and this derivative is positive for $v < u$ but negative for $v > u$. Hence $c_\lambda(u,v)$ as function of v has a single maximum at $v = u$ and is thus a suitable objective function for the purpose of achieving $v = u$ by gradient ascent in v.

In terms of $c_\lambda(u,v)$ the instantaneous cost function for our plasticity rule is

$$c(g_{\mathrm{E}}, g_{\mathrm{I}}; V^*_{\mathbf{w}}) = c_\lambda(U_{\mathrm{M}}, V^*_{\mathbf{w}}) \tag{S6}$$

with λ given by Eq. (S3) and U_{M} given in the main text (Eq. 3). In view of (S5) the partial derivative with respect to the strength of the i-th synapse is

$$\tfrac{\partial}{\partial w_i} c(g_{\mathrm{E}}, g_{\mathrm{I}}; V^*_{\mathbf{w}}) = h(V^*_{\mathbf{w}}) \left(\phi\left((1-\lambda)U_{\mathrm{M}} + \lambda V^*_{\mathbf{w}} \right) - \phi(V^*_{\mathbf{w}}) \right) \tfrac{\partial}{\partial w_i} V^*_{\mathbf{w}} .$$

Using (S3) this can be rewritten as

$$\tfrac{\partial}{\partial w_i} c(g_{\mathrm{E}}, g_{\mathrm{I}}; V^*_{\mathbf{w}}) = \left(\phi\left(U\right) - \phi(V^*_{\mathbf{w}}) \right) h(V^*_{\mathbf{w}}) \tfrac{\partial}{\partial w_i} V_{\mathbf{w}} + \mathcal{O}(\tfrac{1}{g_{\mathrm{D}}}) \tag{S7}$$

Upto the $\frac{1}{g_{\mathrm{D}}}$ correction, the above right hand side is equal to the expectation value of our plasticity rule (Eq. 4, main text). Hence the proposed plasticity rule implements a stochastic gradient ascent procedure on the temporal average of (S6), driving the weight vector $\mathbf{w}$ towards a value satisfying $V^*_{\mathbf{w}} = U_{\mathrm{M}}$.

2.2 Mean field analysis of unsupervised learning

Considering a network of N compartmental neurons with somato-somatic connection matrices A (excitation) and B (inhibition), we derive a self-consistency relation for the vector $\mathbf{R}_{\mathrm{M}}$ of the neuronal firing rates after learning. Our starting point is the equation for the matching potential in a single neuron (3, main text), which for convenience we state in the vector form

$$\mathbf{U}_{\mathrm{M}}(t) = \frac{\mathbf{g}_{\mathrm{E}}(t) E_{\mathrm{E}} + \mathbf{g}_{\mathrm{I}}(t) E_{\mathrm{I}}}{\mathbf{g}_{\mathrm{E}}(t) + \mathbf{g}_{\mathrm{I}}(t)} . \tag{S8}$$

Here $\mathbf{U}_{\mathrm{M}}, \mathbf{g}_{\mathrm{E}}$ and $\mathbf{g}_{\mathrm{I}}$ denote the vectors of matching potential, excitatory nudging and inhibitory nudging for the N neurons. The division in (S8) is component-wise.

If N and/or the firing rates are large enough, we can neglect the temporal fluctuations due to the Poisson spiking. Then the rate vector $\mathbf{R}_{\mathrm{M}}$ determines all of the quantities in the above equation. In particular

$$\begin{aligned}\mathbf{R}_{\mathrm{M}} &= \phi(\mathbf{U}_{\mathrm{M}}) \\ \mathbf{g}_{\mathrm{E}} &= \alpha\, \mathsf{A}\, \mathbf{R}_{\mathrm{M}} \\ \mathbf{g}_{\mathrm{I}} &= \alpha\, \mathsf{B}\, \mathbf{R}_{\mathrm{M}}\end{aligned}$$

where the scaling factor α accounts for the synaptic release kernel. For simplicity, we have assumed the kernel to be the same for excitation and inhibition. Plugging this into Eq. (S8) yields the desired self-consistency relation:

$$\mathbf{R}_{\mathrm{M}} = \phi\left(\frac{(E_{\mathrm{E}}\, \mathsf{A} + E_{\mathrm{I}}\, \mathsf{B})\mathbf{R}_{\mathrm{M}}}{(\,\mathsf{A} + \,\mathsf{B})\mathbf{R}_{\mathrm{M}}}\right). \tag{S9}$$

In general this equation will have many solutions. For instance, if A and B have topographic structure and a specific vector $\mathbf{R}_{\mathrm{M}}$ satisfies (S9), a topographically shifted copy of the vector will also satisfy the equation. One important property of the learning dynamics can however be read of directly from (S9), because the equation is invariant under common rescalings of A and B, i.e under the transformation $\mathsf{A} \to \gamma\, \mathsf{A}$, $\mathsf{B} \to \gamma\, \mathsf{B}$. So the set of possible outcomes of learning is determined only by the balance of somatic excitation and inhibition but not by the absolute strength of the nudging.

3 Simulation Details

The differential equations were integrated using Euler's method with a time step of 0.2. Initial values of the weights of dendritic synapses were picked from a Gaussian distribution with mean μ and standard deviation 2μ, the value of μ being task dependent (see below).

3.1 KL divergence for firing rates

To evaluate learning performance (e.g. in Fig. 1c), we want to asses how close spike trains produced by a neuron with somatic potential $U(t)$ are to the spike trains which would be produced if the potential where $U_{\mathrm{M}}(t)$. Neglecting refractoriness, we assume that the potentials give rise to Poisson spike trains

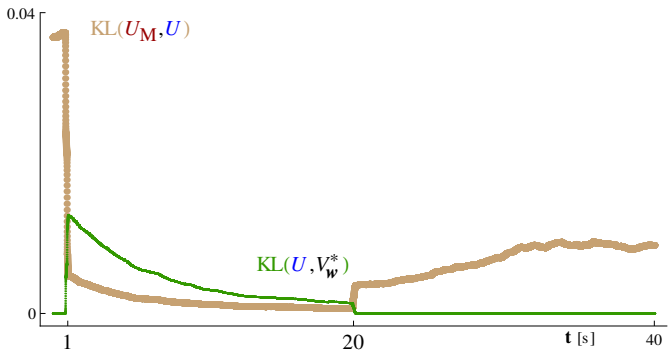

Fig. S5. Evaluation of stability in the absence of nudging. The plot is analogous to Fig. 1c of the main text. But the time frame for which the simulation was run after learning, with ongoing plasticity in the absence of any direct synaptic input to the soma, was extended from 4s in the main text to the 20s shown above.

with rates $\phi(U(t))$ and, respectively, $\phi(U_{\rm M}(t))$. In a short time bin of duration δ the firing probabilities then are $q\delta$ and $p\delta$ for

$$q = \phi(U(t)) \quad \text{and} \quad p = \phi(U_{\rm M}(t))\,. \tag{S10}$$

The firings in the time bin are Bernoulli random variables and the textbook definition for their KL-divergence yields:

$$\mathrm{kl}(U_{\rm M}(t), U(t)) = p\delta \log \tfrac{p\delta}{q\delta} + (1-p\delta) \log \tfrac{1-p\delta}{1-q\delta}\,.$$

Expanding for small δ this becomes

$$\mathrm{kl}(U_{\rm M}(t), U(t)) = \Big(p \log \tfrac{p}{q} + q - p\Big)\,\delta + \mathcal{O}(\delta^2) \tag{S11}$$

Assuming the stimulus is presented from $t = 0$ to $t = T$, averaging over the stimulus duration and using (S10,S11), we set:

$$\mathrm{KL}(U_{\rm M}, U) = \tfrac{1}{T} \int_0^T \mathrm{d}t\ \phi(U_{\rm M}(t)) \log \tfrac{\phi(U_{\rm M}(t))}{\phi(U(t))} + \phi(U(t)) - \phi(U_{\rm M}(t))\,. \tag{S12}$$

The analogous formula is used for $\mathrm{KL}(U, V^*_{\mathbf{w}})$, when comparing the actual somatic potential to the dendritic prediction $V^*_{\mathbf{w}}$.

3.2 Details for Figure 1, main text

The dendritic input pattern shown in Fig. 1a is made up of 200 homogeneous Poisson spike trains, each with a mean rate of 10 HZ. The mean value of a dendritic weight was $\mu = 0.2$ before learning, and a learning rate of $\eta = 0.07$ was used.

We also evaluated in more detail the effects of the synaptic diffusion occurring in the absence of nudging as a consequence of the stochastic nature of the plasticity rule (Fig. S5). While the diffusion is detrimental, performance is still much better than before learning even after 20s of plasticity in the absence of nudging

3.3 Details for Figure 2, main text

Learning time was divided in epochs with a duration of 500 ± 100 ms (Gaussian distribution). For each epoch one of the nudging patterns in Fig. 2a was picked at random and applied during the entire epoch. Total learning time elapsed between Fig. 2b and Fig. 2c was 500 s (simulated biological time). The mean value of a dendritic weight was $\mu = 0.1$ before learning, and a learning rate of $\eta = 0.01$ was used.

We quantitatively evaluated network performance using a protocol similar to the one shown in Fig. 2b and 2c. For each trial a randomly chosen pattern was nudged for 50 ms, followed by a 100 ms period without any nudging before the onset of the next trial. During the 100 ms without nudging we evaluated the persistence of the pattern which had last been nudged by the average value of $\mathrm{KL}(U_{\mathrm{M}}, U)$. The average is over the neurons making up the last pattern as well as over the 100 ms duration. For each neuron and time point the appropriate value of U_{M} for this pattern, obtained from the g_{E} values in Fig. 2a, was used in computing $\mathrm{KL}(U_{\mathrm{M}}, U)$. Before learning, the average value of this discrepancy measure was 0.065 ± 0.001. During the 500 s of learning considered in the main text this measure decreased five-fold to 0.013 ± 0.002.

3.4 Details for Figure 3, main text

The inhibitory neurons in the network were point neurons with a soma modeled like the soma of the compartmental neurons (Eq. 1, main text; Eq. 10, Methods), but for the fact that $g_{\mathrm{D}} = 0$ in the case of the inhibitory neurons. Interaction between compartmental and inhibitory neurons was conductance based and modeled just like the interaction between the somata of compartmental neurons (Eq. 9, Methods).

For the topographic connectivity both the 40×40 compartmental neurons and the 20×20 were placed on square grids within the unit square. To compensate for the relatively small number of neurons we are able to simulate, periodic boundary conditions were used. For implementing these, we define the cyclic

distance of two points p and q in the unit square as

$$d_{\rm c}(p,q) = \sqrt{|p_1 - q_1|_{\rm c}^2 + |p_2 - q_2|_{\rm c}^2}$$

with $|.|_{\rm c}$ defined as

$$|\delta|_{\rm c} = \min\{\, \mathrm{mod}(\delta,1),\, 1 - \mathrm{mod}(\delta,1)\,\}.$$

The definition of $|.|_{\rm c}$ reflects the fact that the maximal distance between two points on the unit line is 0.5, if one thinks of the endpoints of the line as being one and the same point.

Now, the wiring of the network is given by the following rules:

- If *compartmental* neuron a lies at distance $d_{\rm c}$ from a second *compartmental* neuron b, the probability of there being an excitatory connection from a targeting the soma of b is $p = e^{-10 d_{\rm c}}$. For such a connection, the excitatory synapse conveys conductance with strength $w^{\rm E} = 0.06$.
- If *compartmental* neuron a lies at distance $d_{\rm c}$ from *inhibitory* neuron i, the probability of there being an excitatory connection from a targeting i is $p = e^{-8 d_{\rm c}}$. These connections are facilitating, see Eqs. (S13,S14) below, with a baseline conductance of $w^{\rm E}_{\rm base} = 0.008$
- If *inhibitory* neuron i lies at distance $d_{\rm c}$ from *compartmental* neuron b, the probability of there being a inhibitory connection from i targeting the soma of b is $p = \frac{3}{3+(9 d_{\rm c})^2}$. For such a connection, the inhibitory synapse conveys conductance with strength $w^{\rm I} = 0.3$.
- All of the 100 input neurons project onto the dendrites of all of the compartmental neurons.

The above wiring implements long range inhibition because the distribution of connections from inhibitory neurons to compartmental neurons is scale free. Nevertheless, this distribution peaks at short distances, so there is also substantial local inhibition, competing with the short range excitation. If there is local excitation only, all input patterns tend to get mapped to one and the same area by learning. The reason is that once the lengths of the dendritic weight vectors of the neurons in one area of the map start to increase, this becomes self-reinforcing until all input patterns are mapped to this area. The standard remedy for this is, to use an update rule which does not change the length of a neurons weight vector (Ref. 15, main text). We did not adopt this solution, since we are not aware of a mechanism which might coordinate updates in all the synapses of a neuron so that total synaptic strength remains unchanged. Instead, we found that introducing a competition between short range excitation and inhibition, provides a remedy. In the conductance based formulation, the effective strength of inhibition increases with increasing somatic potential. This dampens the self-reinforcing growth of the weight vectors in an area of the map giving other areas a chance to catch up.

Robustness of map formation can be further improved by shifting the balance of excitatory and inhibitory nudging towards inhibition when there is prolonged high activity in one area of the map. The reason is that the self-reinforcing growth of weight vectors decreases the stimulus selectivity of a neuron's response, and neurons which fire in response to many of the stimuli will show prolonged times of high activity. Shifting the balance can be achieved, for instance, by short term depression in the excitatory somato-somatic connection. Other mechanisms could be short term facilitation in synapses targeting the inhibitory neurons, or even a small after-depolarizing current in these neurons.

For the simulations, we implemented facilitation in the afferents of the inhibitory neurons. So the actual conductance w^{E} of a synapse targeting such a neuron can be higher than its baseline value $w^{\mathrm{E}}_{\mathrm{base}}$ given above. It is obtained as

$$w^{\mathrm{E}} = w^{\mathrm{E}}_{\mathrm{base}}(1+f) \tag{S13}$$

where the facilitation variable f increases with the presynaptic firing rate. In particular,

$$\tau_{\mathrm{f}}\dot{f} = -f + F\sum_{s\in X^{\mathrm{E}}}\delta(t-s) \tag{S14}$$

Here τ_{f} is the time constant for facilitation (we used 200 ms), F determines the facilitation strength (we used $F = 40$), and X^{E} denotes the presynaptic spike train of the synapse. The parameter choices mean that a sustained presynaptic rate of 0.1 kHz leads to a value of f which is approximately 4, i.e to a roughly five-fold increase in synaptic strength.

Learning time was divided into epochs of 100 ms. For each epoch one of the six input patterns in Fig. 3b was chosen at random, and during the epoch the input neurons emitted independent Poisson spike trains with mean frequencies given by the chosen pattern. Total learning time elapsed between the top and bottom row of Fig. 3c was 1600 s (simulated biological time). The mean value of a dendritic weight was $\mu = 0.3$ before learning, and a learning rate of $\eta = 0.015$ was used.