

RESEARCH ARTICLE

Prospective Coding by Spiking Neurons

Johanni Brea^{1,2*}, Alexis Tamás Gaál^{1,3}, Robert Urbanczik^{1†}, Walter Senn^{1,4}

1 Department of Physiology, University of Bern, Bern, Switzerland, **2** School of Computer and Communication Sciences and School of Life Sciences, Brain-Mind Institute, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, **3** Courant Institute of Mathematical Sciences, New York University, New York, New York, United States of America, **4** Center for Cognition, Learning and Memory, University of Bern, Bern, Switzerland

† Deceased.

* johanni.brea@epfl.ch
 CrossMark
 click for updates

Abstract

Animals learn to make predictions, such as associating the sound of a bell with upcoming feeding or predicting a movement that a motor command is eliciting. How predictions are realized on the neuronal level and what plasticity rule underlies their learning is not well understood. Here we propose a biologically plausible synaptic plasticity rule to learn predictions on a single neuron level on a timescale of seconds. The learning rule allows a spiking two-compartment neuron to match its current firing rate to its own expected future discounted firing rate. For instance, if an originally neutral event is repeatedly followed by an event that elevates the firing rate of a neuron, the originally neutral event will eventually also elevate the neuron's firing rate. The plasticity rule is a form of spike timing dependent plasticity in which a presynaptic spike followed by a postsynaptic spike leads to potentiation. Even if the plasticity window has a width of 20 milliseconds, associations on the time scale of seconds can be learned. We illustrate prospective coding with three examples: learning to predict a time varying input, learning to predict the next stimulus in a delayed paired-associate task and learning with a recurrent network to reproduce a temporally compressed version of a sequence. We discuss the potential role of the learning mechanism in classical trace conditioning. In the special case that the signal to be predicted encodes reward, the neuron learns to predict the discounted future reward and learning is closely related to the temporal difference learning algorithm TD(λ).


OPEN ACCESS

Citation: Brea J, Gaál AT, Urbanczik R, Senn W (2016) Prospective Coding by Spiking Neurons. *PLoS Comput Biol* 12(6): e1005003. doi:10.1371/journal.pcbi.1005003

Editor: Peter E. Latham, UCL, UNITED KINGDOM

Received: June 25, 2015

Accepted: June 1, 2016

Published: June 24, 2016

Copyright: © 2016 Brea et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper. Additionally the source code of the simulations is publicly available at <https://github.com/jbrea/prospectiveCoding>.

Funding: The work has been supported by the Swiss National Science Foundation www.snf.ch (personal grant no. 310030L-156863 of WS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Author Summary

Sensory inputs are often predictable. Lightning is followed by thunder, a falling object causes noise when hitting the ground, our skin gets wet when we jump into the water. Humans learn regularities like these without effort. Learned predictions allow to cover the ears in anticipation of thunder or close the eyes just before an object hits the ground and breaks into pieces. What changes in the brain when new predictions are learned? In this article, we present a mathematical model and computer simulations of the idea that the activity of a single neuron represents expected future events. Such a prospective coding can be learned in a neuron that receives input from the memory trace of a first event (e.g.

lightning) and also input from the second event (e.g. thunder). Synaptic input connections from the memory trace are potentiated such that the spiking activity ramps up towards the onset of the second event. This deviates from the classical Hebbian learning that merely associates two events that are coincident in time. Learning in our model associates a current event to future events.

Introduction

Animals can learn to predict upcoming stimuli. In delayed paired-associate tasks, animals learn to respond to pairs of stimuli (e.g. images A1-B1 and A2-B2) separated by a delay. These tasks can be solved by either keeping a memory of the first stimulus (A1 or A2) during the delay period (retrospective coding) or anticipating the second stimulus (B1 or B2) during the delay period (prospective coding). Monkeys seem to use both coding schemes [1]. Recordings in the prefrontal cortex of monkeys performing a delayed paired-associate task revealed single neurons with decreasing firing rate in response to a specific first stimulus (A1 or A2) and other neurons with ramping activity in trials where a specific second stimulus (B1 or B2) is anticipated [1, 2]. Thus, the firing rate of a neuron may encode not only past and current events, but also prospective events.

Learning to anticipate a future stimulus can also be observed in classical trace conditioning, where a conditioned stimulus (CS, e.g. sound of a bell) is followed after a delay by an unconditioned stimulus US (e.g. a sausage) that causes a response R (e.g. salivation) [3, 4]. After several repetitions of this protocol, the conditioned stimulus CS can elicit response R already before the onset of the unconditioned stimulus US.

A common experimental finding in these examples is the slowly ramping neuronal activity prior to the predicted event. In an experiment where mice choose to lick left or right in response to a tactile cue, the neural activity in the anterior lateral motor cortex ramps up in the waiting period before the response [5]. This activity pattern implements prospective coding as it indicates whether the animal will lick left or right. Serotonergic neurons in the dorsal raphe nucleus of mice show an activity ramp in a delay period between a predictive odor cue and the availability of a sucrose reward [6]. In rats that navigate a maze towards the learned position of a chocolate milk reward, the activity of striatal neurons increases while the rat approaches the reward position [7, 8]. In visual delayed paired associate tasks in which monkeys are trained to select a specific choice object that is associated with a previously shown cue object, increasing activity in the delay period was measured for neurons in the prefrontal cortex [1, 9, 10] and in the inferior temporal cortex [2, 11].

It is unclear how prospective coding emerges. The cue and the associated predictable event are typically separated by an interval of some seconds. On the other hand, synaptic plasticity, that is presumably involved in learning new associations, typically requires presynaptic and postsynaptic activity to coincide in a much shorter interval. Some tens of milliseconds is, for example, the size of the ‘plasticity window’ in spike-timing dependent plasticity; no synaptic change occurs, if presynaptic and postsynaptic spike are separated by more than the size of this plasticity window [12, 13]. This mismatch between the behavioral and the neuronal timescales begs the question how a neuronal system can learn to make predictions more than a second ahead. There are also plasticity mechanisms that can correlate pre- and postsynaptic spiking events that are separated by seconds [14, 15]. Yet, assuming many simultaneously active afferents, it remains unclear how the behaviourally relevant pair of pre- and postsynaptic spikes can be selected out of hundreds behaviourally irrelevant pairs.

In normative models of synaptic plasticity, the shape of the causal part of the plasticity window matches the shape of the postsynaptic potential (PSP), if the objective is to reproduce precise spike timings [16–18]. However, if the objective is to reproduce future activity, this specific learning rule is insufficient. Yet, as we demonstrate in this article, the same plasticity rule with only a slightly wider window also allows for learning a prospective code. With this mechanism, it is possible to learn an activity ramp towards a specific event in time, or to learn predicting a time-varying signal or a sequence of activities well ahead in time. In a 2-compartment neuron model, this mechanism leads to the dendritic prediction of *future* somatic spiking. The mechanism stands in contrast to the work of Urbanczik & Senn, where the current somatic spiking is predicted [18]. Despite this fundamental difference, the plasticity rules only differ in the width of the potentiation part of the plasticity window.

Results

Schematic description of the learning mechanism

Before defining the learning rule in detail, we provide an intuitive description. In a neuron with both static synapses (green connection in Fig 1A and 1B) and plastic synapses (blue in Fig 1A and 1B), we propose a learning mechanism for the plastic synapses that relies on two basic ingredients: spike-timing dependent synaptic potentiation and balancing synaptic depression. The synaptic connections are strengthened if a presynaptic spike is followed by a postsynaptic spike within a ‘plasticity window of potentiation’ (red in Fig 1A and 1B). The size of this plasticity window turns out to have a strong influence on the timing of spikes that are caused by strengthened dendritic synapses. If the plasticity window has the same shape as a postsynaptic potential (PSP), learned spikes are fired at roughly the same time as target spikes [16–18]. But if the plasticity window is slightly longer than the postsynaptic potential, learned spikes tend to be fired earlier than target spikes. More precisely, because of the slightly wider plasticity window of potentiation, presynaptic spikes may elicit postsynaptic spikes through newly strengthened connections (thick blue arrow in Fig 1B) even before the onset of the input through static synapses. These earlier postsynaptic spikes allow to strengthen the input of presynaptic neurons that spike even earlier. We refer to this as the bootstrapping effect of predicting the own predictions. As a result, a postsynaptic activity induced by the input through static synapses will be preceded by an activity ramp produced by appropriately tuned dendritic input. The neuron learns a prospective code that predicts an upcoming event.

The 2-compartment neuron model

We consider a 2-compartment neuron model that captures important functional details of spiking neurons and is well suited for analytical analysis [18]. In this model (Fig 1C), a dendritic compartment receives input through plastic synapses with strength w . The voltage U of the somatic compartment is coupled to the dendritic voltage V_w and receives additional input I through static synapses,

$$C \dot{U} = -g_L U + g_D (V_w - U) + I, \tag{1}$$

where g_L is the leak conductance, g_D is the coupling conductance between soma and dendrite and C , the somatic capacitance. The dendritic potential V_w is given by a weighted sum of presynaptic inputs, i.e.

$$V_w(t) = \sum_i w_i \text{PSP}_i(t) = \sum_i w_i \sum_{t^f \in \mathcal{T}_i} \kappa(t - t^f) \tag{2}$$

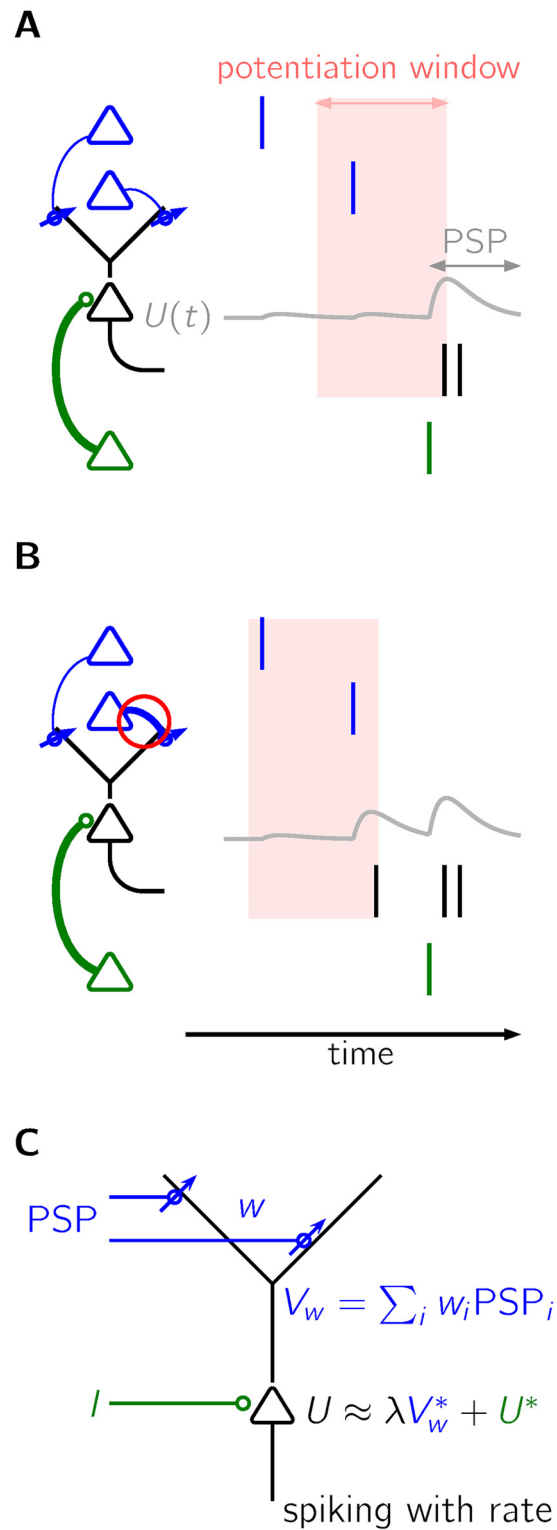


Fig 1. A learning mechanism that leads to prospective coding. **A** The signal to be predicted (target input) originates from the green neuron and depolarizes the black neuron (gray trace) such that it spikes (black lines). The synaptic connection between a blue neuron and the black neuron is strengthened if pre- and postsynaptic spikes lie within the red plasticity window of potentiation, which is slightly broader than a typical postsynaptic potential. **B** Due to the strengthened connection (red circle), the black neuron spikes already

before the target input arrives. Since earlier presynaptic spikes now also lie within the potentiating plasticity window, the activity of the black neuron will be anticipated earlier, giving rise to prospective coding. **C** A spiking neuron receives input through plastic dendritic synapses with strengths w_i and an input I through static (i.e. non-plastic) synapses. The somatic membrane potential U is well approximated by the sum of attenuated dendritic input V_w^* and attenuated somatic input U^* .

doi:10.1371/journal.pcbi.1005003.g001

with plastic synaptic weights w_i , postsynaptic potentials PSP_i that model the depolarization of the postsynaptic membrane potential due to the arrival of a presynaptic spikes at synapse i , set \mathcal{T}_i of spike arrival times at synapse i and spike response kernel κ . Spiking of the postsynaptic neuron is modeled as an inhomogeneous Poisson process with rate $\varphi(U)$.

We model the input with time varying excitatory and inhibitory conductances g_E and g_I proximal to the soma such that

$$I(t) = g_E(t)(E_E - U) + g_I(t)(E_I - U) \quad (3)$$

as proposed by Urbanczik & Senn [18].

For large total conductance and slowly varying input, the somatic membrane potential $U(t)$ is well approximated (see [Methods](#)) by its steady state solution

$$U(t) \approx \lambda(t)V_w^*(t) + U^*(t), \quad (4)$$

where we introduced the attenuated dendritic potential

$$V_w^*(t) = \frac{g_D}{g_L + g_D} V_w(t), \quad (5)$$

the attenuated somatic input

$$U^*(t) = \frac{g_E(t)E_E + g_I(t)E_I}{g_{\text{tot}}(t)} \quad (6)$$

and the ‘nudging’ factor

$$\lambda(t) = \frac{g_L + g_D}{g_{\text{tot}}(t)}, \quad (7)$$

with $g_{\text{tot}}(t) = g_L + g_D + g_E(t) + g_I(t)$, to be in accordance with Urbanczik & Senn [18]. The nudging factor $\lambda(t) \in (0, 1]$ is close to 1 for small somatic input and equal to 1 if $g_E(t) + g_I(t) = 0$.

Learning as dendritic prediction of the neuron’s future discounted firing rate

The plasticity rule we consider for the dendritic synapses can be seen as differential Hebbian in the sense that both the potentiation and depression term are a product of a post- and presynaptic term. The strength of synapse i is assumed to change continuously according to the dynamics

$$\dot{w}_i = \eta \left(\alpha \varphi(U) \widetilde{PSP}_i - \varphi(V_w^*) PSP_i \right), \quad (8)$$

where

$$\widetilde{PSP}_i(t) = \frac{1}{\tau} \int_0^\infty ds e^{-\frac{t-s}{\tau}} PSP_i(t-s) \quad (9)$$

is the low-pass filtered postsynaptic potential at synapse i , $\varphi(U)$ and $\varphi(V_w^*)$ are the

instantaneous firing rates based on the somatic potential and the attenuated dendritic potential, respectively, and η is the learning rate. The factor of potentiation α that scales the potentiation term is positive but smaller than the inverse of the largest nudging factor $1/\max_t \lambda(t)$ to prevent the unbounded growth of synaptic strengths.

Under the assumption of a periodic environment, rich dendritic input dynamics, constant nudging factor λ and linear φ (Methods), the weight dynamics in Eq 8 leads to prospective coding by making the dendritic rate $\varphi(V_w^*(t))$ approach the expected future discounted somatic input rate, i.e.

$$\varphi(V_w^*(t)) = \frac{\alpha}{\tau} \int_0^\infty ds e^{-\frac{s}{\tau_{\text{eff}}}} \varphi(U^*(t+s)), \tag{10}$$

where the effective discount time constant τ_{eff} is given by

$$\tau_{\text{eff}} = \frac{\tau}{1 - \lambda\alpha}. \tag{11}$$

Depending on the factor of potentiation α and the nudging factor λ , the effective time constant τ_{eff} can be much larger than the biophysical time constant τ of low-pass filtering and match behavioral timescales of seconds. In particular, if the somatic input is strong and hence λ close to 0 (close to ‘clamping’), the effective discount time constant is short, $\tau_{\text{eff}} \approx \tau$. But when nudging is weak (λ close to 1), the synapses on the dendrite learn to predict their self-generated somatic firing rate and the effective discount time constant is extended up to $\tau_{\text{eff}} \approx \frac{\tau}{1-\alpha}$. The case of weak nudging is also the case when the neuron’s somatic firing rate is roughly determined by the dendritic input, $\varphi(U(t)) \approx \varphi(V_w^*(t))$, see Eq 4. In particular, if after learning the somatic input is transiently silenced, the neuron’s firing rate $\varphi(U(t))$, according to Eq 10, represents the discounted future rate of the somatic input $U^*(t)$ applied during the previous learning period, even if this was only slightly nudging the somatic potential $U(t)$ itself.

Periodic inputs are unrealistic in a natural setting. But a similar result holds also in more general settings, where a neuron is occasionally exposed to correlated dendritic and somatic inputs. In this more general stochastic setting we derive the main result under the assumption that dendritic and somatic inputs depend on the state of a stationary latent Markov chain X_0, X_1, \dots . The dependence on a stationary latent Markov chain assures that the neuron is occasionally exposed to correlated dendritic and somatic inputs. The main result in this setting is (cf. Eq 48)

$$\varphi(V_w^*(x)) = \frac{\alpha}{1 - \lambda\alpha} \sum_{k=0}^\infty \gamma_{\text{eff}}^k \mathbb{E}[\varphi(U^*(X_k)) | X_0 = x], \tag{12}$$

where $\gamma_{\text{eff}} = e^{-\frac{\delta}{\tau_{\text{eff}}}}$ is a large discount factor that leads to a similar discount behavior as in the time-continuous case, if $t = k\delta$.

It is important to note that in the stochastic case the dendritic rate is only informative about *expected* future somatic inputs. Metaphorically speaking, a neuron can learn to predict the expected win in a lottery, but obviously it cannot learn to predict single lottery draws.

The bootstrapping effect of predicting the own predictions

In the limit, $\tau \rightarrow 0$ we find that $\widetilde{\text{PSP}} = \text{PSP}$ and with $\alpha = 1$ we recover the learning rule of Urbanczik & Senn [18]. This rule adapts the dendritic synapses such that the dendritic input matches the somatic input Fig 2B. On the other hand, the learning rule with a slightly larger potentiation window leads to dendritic input that ramps up long before the onset of somatic input Fig 2C.

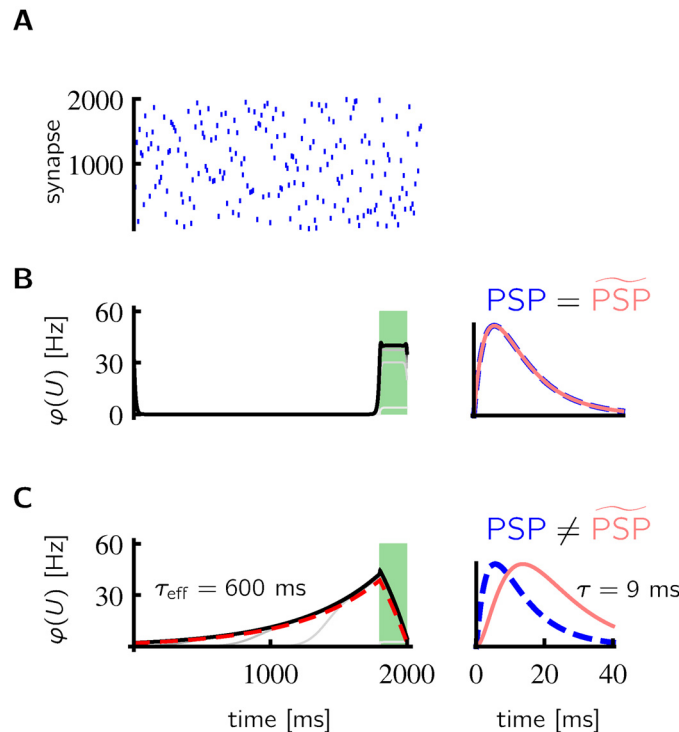


Fig 2. Learning activity ramps on a behavioral timescale with a biologically plausible plasticity window. **A-B** For orthogonal input patterns (exactly one presynaptic spike arrives at each synapse during 2 s) (A) and a somatic input after 1800 ms ($g_E = 15$ nS during green shading in B and C, $g_E = 0$ otherwise), the learned postsynaptic firing rate has a similar time course as the somatic input if $PSP = \widetilde{PSP}$ (B, lines from light gray to black: postsynaptic firing rate after 100, . . . , 1000 training sessions). **C** If $PSP \neq \widetilde{PSP}$ (with $\tau = 9$ ms), the learned postsynaptic firing rate ramps up with an effective time constant of $\tau_{\text{eff}} = 600$ ms towards the onset of the somatic input. The theoretical result is in good agreement with the simulation (dashed red line: $\varphi(V_w^*)$ computed by Eq 10). During training, the 2 s long pattern of dendritic and the somatic inputs is periodically repeated.

doi:10.1371/journal.pcbi.1005003.g002

By looking at Eqs 4 and 8 we can now obtain a better intuition for the bootstrapping effect of predicting the own predictions. If at the beginning of learning all synaptic weights w_i are zero, the dendritic potential V_w is at rest (= 0) all the time and the somatic membrane potential $U(t)$ follows the somatic input $U^*(t)$ (see Eq 4). In this case, the learning rule in Eq 8 contains only the potentiation term

$$\dot{w} = \eta \alpha \varphi(U^*) \widetilde{PSP}_i. \quad (13)$$

In the example in Fig 2C, the somatic input U^* and consequently $\varphi(U^*)$ is non-zero only after 1800 ms. Therefore, synapse i is potentiated only if a presynaptic spike arrives shortly before the onset of the somatic input. The next time a presynaptic spike arrives at synapse i , the somatic membrane potential is depolarized by the dendritic input already before the onset of the somatic input and the learning rule contains at this moment (e.g. at 1780 ms in Fig 2C) the terms

$$\dot{w} = \eta \left(\alpha \varphi(\lambda V_w^*) \widetilde{PSP}_i - \varphi(V_w^*) PSP_i \right). \quad (14)$$

These terms would cancel each other in the case of Urbanczik & Senn [18] where $\alpha = \lambda = 1$ and $\widetilde{PSP}_i = PSP_i$. But if \widetilde{PSP}_i is the low-pass filtered version of the postsynaptic potential (as in Fig 2C) they do not cancel. Instead, synapses are potentiated, if a presynaptic spike arrives

shortly before the somatic potential was depolarized due to dendritic input through already potentiated synapses. The consequence of this bootstrapping effect appears in Fig 2C in the gray curves. After 100 training sessions, the dendritic input starts to rise around 1200 ms, but synapses with earlier presynaptic spikes are not yet strengthened. With each further training session the dendritic input rises earlier.

The dendritic and the somatic inputs are deterministic periodic functions, in the example in Fig 2C. Therefore we can directly compare the simulation to the theoretical results of the previous section. For the interval without somatic input (0–1800 ms), where $\varphi(U) = \varphi(V_w^*)$, we find a good agreement (dashed red and thick black line in Fig 2C). Small differences are to be expected, because in the theoretical derivations a constant nudging factor λ is assumed and the steady-state solution of the somatic membrane potential dynamics is used (see Eq 4). The dendritic rate $\varphi(V_w^*)$ is only slightly below the somatic rate $\varphi(U)$ in the interval with somatic input (1800–2000 ms), because the somatic input is small.

Dependence on the dendritic input structure

The input pattern in Fig 2A is a particularly simple example of a deterministic, periodic pattern with rich enough structure. Enough structure to learn a prospective code exists also in sufficiently many randomly generated (frozen) spike trains that are deterministically repeated, if there is always at least one presynaptic spike within the duration of a PSP and the probability of repeating a nearly identical presynaptic spike pattern is low (see Fig 3A). We did not systematically search for the minimal number of required dendritic synapses. But for the example in Fig 3A we found empirically that a few hundred synapses are necessary. If the presynaptic firing frequency is only 2 Hz, we found that 1000 presynaptic neurons are enough to learn the ramp in 100 trials, whenever the learning rate is larger than in the 20 Hz case. At the end of learning, the time course of the somatic potential matches the one of the previous example (black lines in Figs 2C and 3A). But during learning, the time course of the somatic potential is different in the two examples (gray lines in Figs 2C and 3A). This is a consequence of the influence of correlations in the dendritic input. For the frozen spike trains, the presynaptic auto-correlation $\mathbb{E}[\text{PSP}_i(t)\text{PSP}_i(t+s)] \neq 0$ is non-vanishing for all s and i . This causes the average firing rate to increase early during learning (Fig 3A; gray lines in interval 0–1500 ms in contrast to gray lines in the same interval in Fig 2C).

In the examples given so far, the dendritic and the somatic inputs are deterministic, but deterministic repetitions of the exact same spike trains are unrealistic. In Fig 3B we consider the more realistic case of random spiking. In each trial, the spikes are sampled from an inhomogeneous Poisson process, with periodically repeating rates. The resulting activity ramp is noisier but in good agreement with the theoretical result. It is important that the rates of the Poisson process have sufficiently rich structure. In Fig 3D the firing rate of the Poisson process is kept constant for one third of the trial. In this case, the temporal structure is not sufficiently rich to learn a smooth ramp and a stepwise activity ramp is learned instead.

In Fig 3C, the target event occurs only with a 50% chance, i.e. the somatic input is given only in half the trials. This results in an activity ramp with smaller amplitude, which is consistent with the theoretical finding that the dendritic rate depends linearly on the average somatic input rate (see Eq 12).

Delayed paired-associate task

Prospective coding in neurons of the prefrontal cortex was observed in an experiment with monkeys performing a delayed paired-associate task [1]. In this experiment, monkeys learned to associate a visual sample to a visual target presented one second later. Our learning rule allows for learning a prospective code in such a task.

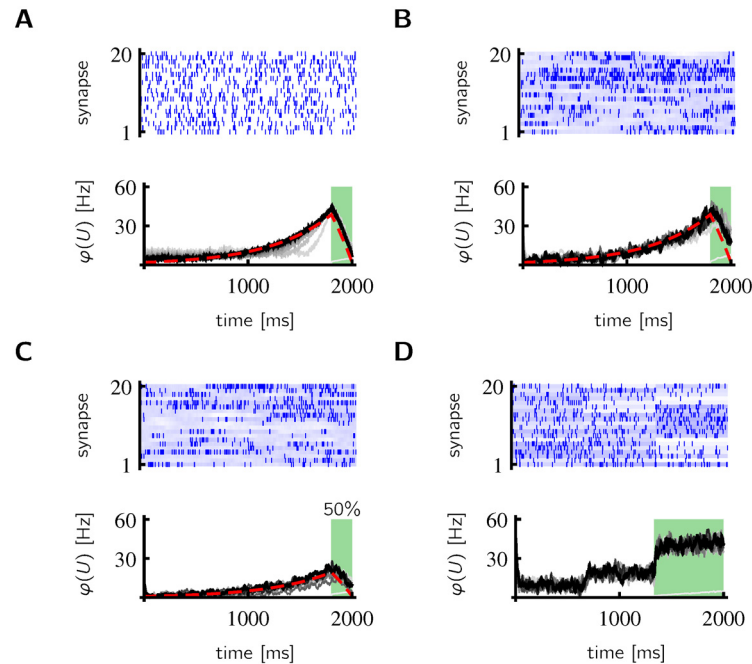


Fig 3. Dependence of learning on the type of input. **A** Learning succeeds with deterministically repeated spike trains (20 frozen spike trains out of 500, g_E as in Fig 2). **B** Learning succeeds with stochastic spiking, if the spiking rate is variable (blue shading: spiking rate, blue ticks: one sample). **C** The amplitude of the ramp is smaller, if the training event occurs only with 50% probability. **D** If the spiking rate is constant during long intervals, the input is not sufficient to learn a smooth ramp. A stepwise ramp is learned instead ($g_E = 8$ nS during green shading)

doi:10.1371/journal.pcbi.1005003.g003

During training, sample A1 is always followed by target B1 after a delay of 1s, and sample A2 is followed by target B2 (Fig 4A). In the simulation we assume that the sample (first stimulus) leaves a memory trace in form of a spatio-temporal activity pattern that projects through dendritic synapses, while the target (second stimulus) drives somatic synapses (Fig 4B). In order to have sufficiently rich presynaptic activity (c.f. Fig 3B), the memory trace of the sample is modeled by an inhomogeneous Poisson process with sample dependent rate trajectories (Fig 4C), i.e. during the presentation of the first stimulus the rate trajectory of each neuron approaches a previously chosen template trajectory that depends on the sample (see Methods). These memory traces are inspired by liquid state machines (see Discussion). If a neuron receives strong somatic input only in the presence of a specific target (neurons 1 and 2 in Fig 4B), its firing rate ramps up exclusively in anticipation of this target (neurons 1 and 2 in Fig 4D). In contrast to such a ‘grandmother-cell coding’ (one neuron for one target), a set of neurons could encode the target in a distributed manner, where the target is identified by the overall activity pattern and single neurons respond differently to different target stimuli. Such a distributed code can be learned with neurons that receive somatic input of target-specific strengths (neuron 3 in Fig 4B; B1 stronger than B2). After learning, the amplitude of the activity ramp reflects this target specificity (neuron 3 in Fig 4D).

Prospective coding of times series

In Figs 2 to 4 the somatic target input was silent most of the time and active only during a short interval. This simple time course of the somatic input is, however, not a requirement and learning also converges for more complex trajectories of somatic input. In general, a time varying

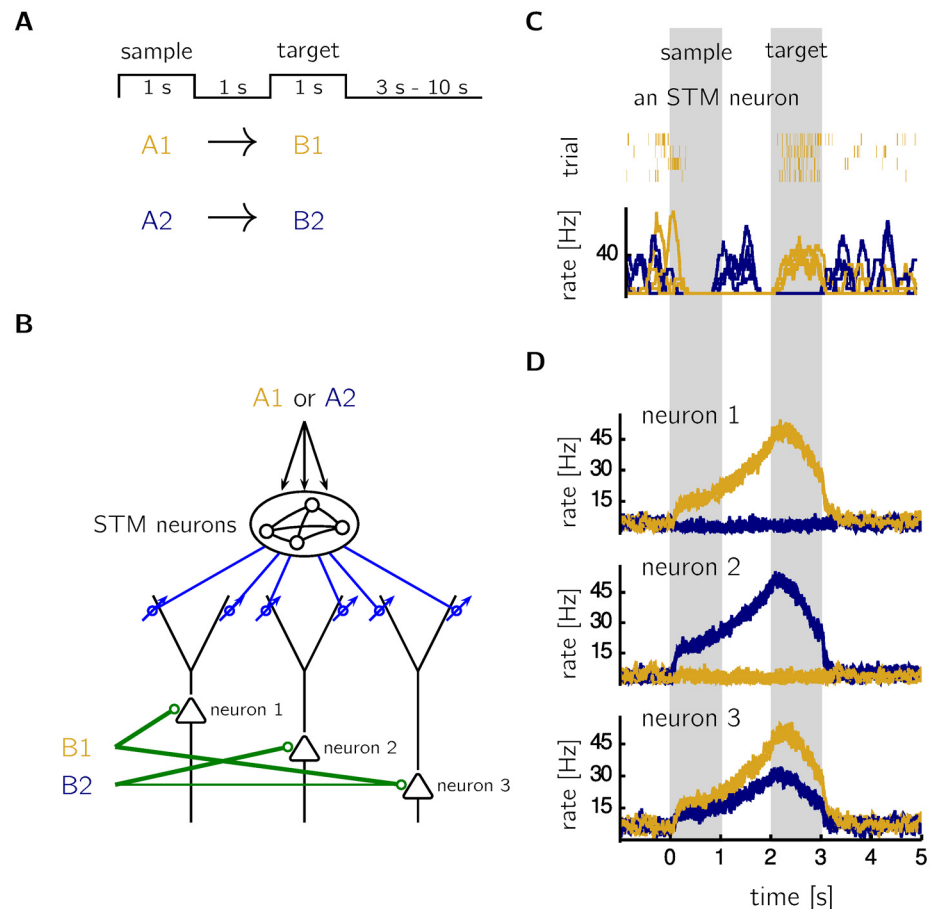


Fig 4. Prospective coding in a delayed paired-associate task. **A** In the simulation, stimulus A1 and A2 is repeatedly followed by stimulus B1 and B2, respectively, with a delay of 1s. The two pairs are chosen randomly with equal probabilities. Intertrial intervals are chosen at random uniformly between 3 s and 10 s. **B** The first stimulus (A1 or A2) activates a recurrent network of 2000 neurons representing a short-term memory (STM). The dynamics of the recurrent network is modeled by a stochastic process (see [Methods](#)). The STM is read out by 3 neurons that encode in a distributed manner the second stimulus (setting $g_E = 5$ nS in neuron 1 and neuron 3 during the B1 presentation, and $g_E = 5$ nS in neuron 2 and $g_E = 2.5$ nS in neuron 3 during the B2 presentation). **C** The time course of the firing rates of neurons in the recurrent short-term memory network depends on the first stimulus (dark blue: A1; gold: A2; spike trains of a specific STM neuron during 4 A1 trials and its estimated rates for 4 A1 and 4 A2 trials). **D** After learning, the firing rate of neuron 1 ramps up after stimulus A1 (gold trace), but not after stimulus A2 (blue trace). The opposite holds for neuron 2. Since neuron 3 receives more somatic input when B1 is present, the firing rate of neuron 3 ramps up to a larger value after A1 than after A2.

doi:10.1371/journal.pcbi.1005003.g004

input through (static) somatic synapses induces plasticity that advances the postsynaptic firing rate $\varphi(U(t))$ relative to the firing rate $\varphi(U^*(t))$ determined by the somatic input alone. [Fig 5A](#) shows an example with an advancement of roughly 50ms that has been achieved with a shorter time window (~ 20 ms) for synaptic potentiation. As in [Fig 3A](#), the dendritic input was a periodically repeated random spike train that could also be replaced by stochastic spiking with time dependent firing rates as in [Fig 3B](#).

Since the learning rule converges to a point where the dendritic input is proportional to the future discounted somatic input ([Eq 10](#)), the advanced sequence (black in [Fig 5A](#)) is not simply a forward shifted version of the somatic input (green in [Fig 5A](#)). This becomes clearly apparent at the center of the figure, where the somatic input is symmetric around 1000 ms, but the

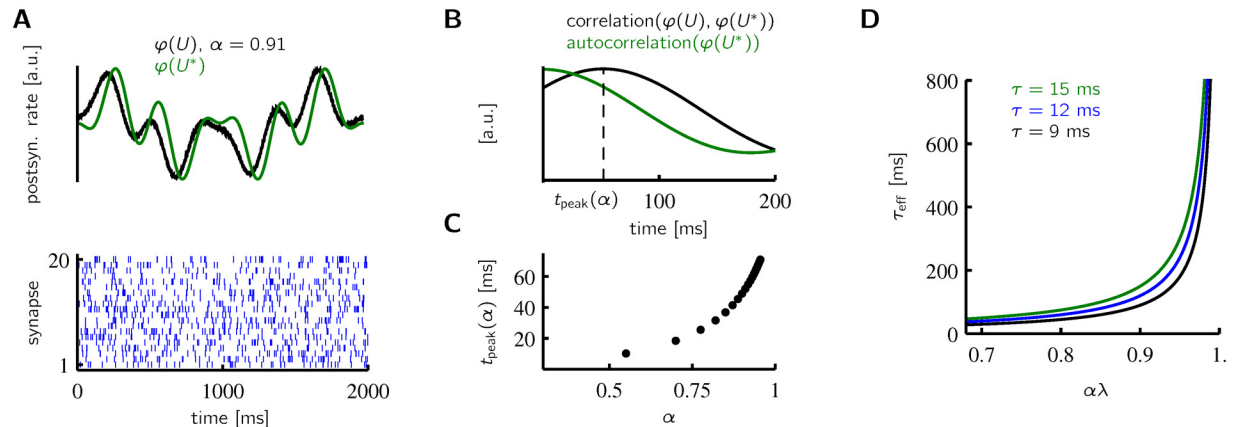


Fig 5. Prospective coding of time-varying input. **A** Learning leads to an advancement of the postsynaptic firing rate. The dendritic input consists of spike trains of 2000 neurons (bottom; 20 shown). The somatic input is given by $g_E(t) = 6(1 - \sin(\omega t) \sin(2\omega t) \cos(4\omega t))$ nS, with $\omega = 2\pi/(2000 \text{ ms})$. **B** The correlation of the firing rate curves in A peaks at $t_{\text{peak}} = 52$ ms. **C** The advancement increases with the potentiation factor α . **D** With increasing potentiation factor $\alpha\lambda$ the effective discount time constant τ_{eff} becomes much larger than τ (Eq 11).

doi:10.1371/journal.pcbi.1005003.g005

advanced sequence is decaying, because the somatic input has a strong dip around 1100 ms. Despite this, the advancement can be characterized by the peak time of the correlation function between $\varphi(U(t))$ and $\varphi(U^*(t))$ that, as the effective discount time constant τ_{eff} , diverges with increasing potentiation factor α (Fig 5B–5D).

Time series prediction is a fundamental operation of the brain that is, for instance, involved in motor planning. In our context, the activity time course that has to be reproduced may be provided by proprioceptive feedback from muscles as somatic input U^* to neurons in the primary motor cortex [19]. This feedback can be weak, delayed and sparse. The dendritic input V^* , in turn, may be conveyed by a higher visual area or a premotor planning area. This dendritic input learns to predict the discounted future firing rate caused by the somatic input, and hence learns to produce the muscle activity that feeds back again as a delayed proprioceptive signal.

Prospective coding in a recurrent neural network

Lastly, we consider a recurrently connected network of 200 neurons that receive external input only at the soma and no external input at the dendrites. The input at the dendrites is given by the output spikes of the network neurons, where we consider all-to-all connectivity (Fig 6A). In contrast to the examples in Figs 2 to 5, there is no external control to assure the richness of the dendritic input and there are no guarantees that learning converges in the sense of Eq 10. Still, we observe the interesting result that learning changes synaptic strengths to allow fast replay of slow experienced sequences.

For sequentially and periodically repeated stimulations on a slow timescale (green shading in Fig 6), the recurrent dendritic connections between subsequently stimulated groups of neurons are strengthened. After 300 repetitions of the same sequence, a brief initial stimulation is sufficient to evoke an activity sequence that has the same ordering as the original sequence (Fig 6B after 2400 ms). However, the replay dynamics can be much faster than the dynamics of the stimulation. Replay depends on the internal dynamics of the network, notably the time constants of the PSP and the membrane time constant. Due to prospective coding, the sequence becomes advanced in time while repeatedly presenting the stimuli, and due to the recurrent connectivity the advanced sequence can be recalled with a brief stimulation of the first group of neurons (Fig 6B). Note that there is no need to explicitly distinguish between a training and

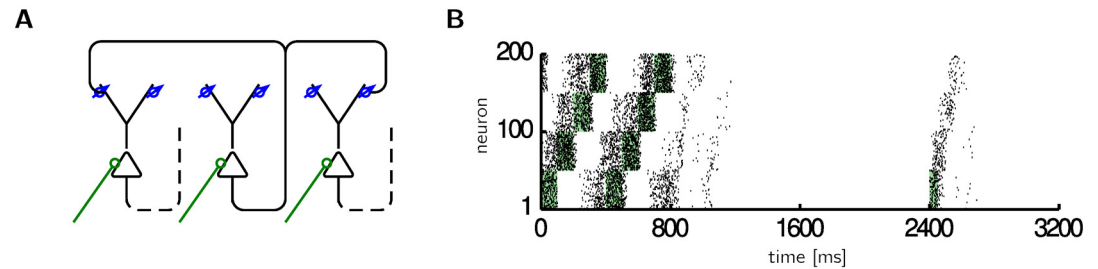


Fig 6. Prospective coding in a recurrent neural network. **A** The dendritic input consists of the spike trains of other neurons within the same network (for clarity only one axon is completely drawn; in the simulation we used all to all connections). **B** Groups of 50 neurons receive sequential somatic input ($g_E = 20$ nS during green shading) of duration 100 ms. After repeatedly stimulating, the firing rate increases already prior to somatic input (see for example neurons 51–100 in the first 100 ms). Before 800 ms the last two of 300 training repetitions are shown. Afterwards, no somatic input is provided anymore except for a brief stimulation (after 2400 ms), after which the sequence is autonomously replayed at a faster speed.

doi:10.1371/journal.pcbi.1005003.g006

recall session. Recall differs from training only in the somatic input, which consists of a brief activation of the first group of neurons during recall and slow, sequential activation during training. The learning rule is active all the time.

Relation to TD(λ)

The proposed learning mechanism of prospective coding is related to a well studied version of temporal difference (TD) learning. Using our notation for a stochastic and time discrete setting, the goal in TD learning is to estimate a value function

$$\mathcal{V}(x) = \frac{\alpha}{1 - \alpha\lambda} \sum_{t=0}^{\infty} \gamma_{\text{TD}}^t \mathbb{E}[\varphi(U^*(X_t)) | X_0 = x], \quad (15)$$

where $0 < \gamma_{\text{TD}} < 1$ is a discount factor and the expectation is taken over the Markov chain X_0, X_1, \dots . We assume that this value function can be approximated by a linear function of the form

$$\widehat{\mathcal{V}}(x) = \varphi(V_w^*(x)), \quad (16)$$

where φ is linear. In TD(λ) with linear function approximation, the weights w evolve according to the learning rule [20–22]

$$\Delta w_{t,i} = w_{t+1,i} - w_{t,i} = \eta \delta_t \widehat{\text{PSP}}_{t,i}, \quad (17)$$

with learning rate η , eligibility trace $\widehat{\text{PSP}}_{t,i} = \sum_{s=0}^{\infty} (\lambda_{\text{TD}} \gamma_{\text{TD}})^s \text{PSP}_i(X_{t-s})$, $0 \leq \lambda_{\text{TD}} \leq 1$, and delta error

$$\delta_t = \frac{\alpha}{1 - \alpha\lambda} \varphi(U^*(X_t)) + \gamma_{\text{TD}} \varphi(V_{w_t}^*(X_{t+1})) - \varphi(V_{w_t}^*(X_t)). \quad (18)$$

This delta error is zero on average if the approximation $\varphi(V_{w_t}^*(x))$ is equal to the value function $\mathcal{V}(x)$ in Eq 15. Furthermore, $\varphi(V_{w_t}^*(x))$ converges to $\mathcal{V}(x)$ under the learning rule of TD(λ) in Eq 17 [20]. The discrete time version of our learning rule (Eq 42), implemented in the 2-compartment model, converges to Eq 12 which is identical to the value function in Eq 15 if $\gamma_{\text{TD}} = \gamma_{\text{eff}}$. Therefore, this form of TD(λ) and our learning mechanism converge to the same value. It is also interesting to see that both methods use an eligibility trace $\widehat{\text{PSP}}$ and $\widetilde{\text{PSP}}$ that

are the same if $\lambda_{TD} \gamma_{TD} = \gamma$, i.e. $\lambda_{TD} = \gamma/\gamma_{eff}$. But despite the convergence to the same point and the use of the same eligibility trace, learning moves in general along different trajectories under this form of TD(λ) and the learning mechanism we propose.

So far we compared the learning mechanism of prospective coding to the plain TD(λ) that has access to the PSP and U^* . If only access to $U = \lambda V^* + U^*$ is available, it is also possible to combine TD(λ) with the bootstrapping effect of predicting the own predictions by implementing a variant of TD(λ) in the dendritic compartment of the 2-compartment model. If the delta error is defined as

$$\delta_t = \alpha \varphi(U(X_t)) + \gamma \varphi(V_{w_i}^*(X_{t+1})) - \varphi(V_{w_i}^*(X_t)), \quad (19)$$

one can show that the learning rule in Eq 8 is almost identical to the TD learning rule in Eq 17 with $\lambda_{TD} = 1$ (Methods). In this case, the weights during learning move along similar trajectories, irrespective of whether this form of TD(1) or our learning rule is used. If this form of TD(1) were not implemented in the 2-compartment model, i.e. if the first term in the delta error in Eq 19 would be replaced by $\varphi(U^*(X_t))$, the time constant of future discounting would be γ instead of γ_{eff} . But since the first term in the delta error in Eq 19 depends on the full somatic potential $U = \lambda V^* + U^*$ the bootstrapping effect of predicting the own predictions applies and the large time constant γ_{eff} arises.

Discussion

As a simple and biologically plausible explanation for how animals can learn to predict future events, we have proposed a local plasticity mechanism that leads to prospective coding in spiking neurons, i.e. the plastic synapses change such that the neuron's current firing rate depends on its expected, future discounted firing rate.

Our model proposes a partial solution to the problem of learning associations on a behavioral timescale without using slow intrinsic processes. Even with a plasticity window that is only slightly larger than the duration of a postsynaptic potential, the effective time constant of discounting the expected future firing rate can be on the order of seconds, thanks to the bootstrapping effect of predicting the own predictions. This effect arises because already predictive inputs influence the activity of a neuron. This is captured by the 2-compartment model of Urbanczik & Senn [18], where the output depends on both the dendritic (student) and the somatic (target) input.

For clarity, we presented the model with target input through static (i.e. non-plastic) somatic synapses and in the examples of ramping activity in Figs 2 and 3 the somatic input was non-zero only during a short period. This simple form of the target input is not a requirement. First, the learning mechanism also applies to arbitrary time courses of the somatic input, as we show in the example of time series prediction in Fig 5, where an advanced and smoothed version of a complex somatic input is learned. Second, the somatic synapses do not need to be static. Yet, they should change slower than the dendritic synapses in order to get a separation of plasticity timescales. And third, the target input could also arrive at another dendritic branch instead of the soma (see generality of the results in Methods).

We focused solely on learning temporal associations and neglected important aspects of learning in animals. However, the proposed learning mechanism can easily be extended to include, for example, a weighting based on behavioral relevance. In the delayed paired-associate task, our model learns the associations between sample and target irrespective of the behavioral relevance of this association. In animal training, however, reward or punishment is crucial; for example the monkeys in the study of Rainer et al. [1] received juice rewards. The

learning rate in our learning mechanism is a free parameter that could incorporate a weighting by behavioral relevance. Biophysically, a neuromodulator like dopamine could implement this modulation of the learning rate. It is also possible to postpone the weight update in Eq 1 and use reward modulated eligibility traces instead (see e.g. [23–25] for theory and [15, 26] for experiments).

The proposed learning mechanism could also be involved in classical trace conditioning, where the first stimulus (CS) is separated from the second stimulus (US) and the response (R) by a delay period, similar to the situation in the delayed paired-associate task. Let us assume that neuron 1 in Fig 4 is involved in initiating response R (e.g. salivation). If the unconditioned stimulus causes somatic input to this neuron and a memory trace of the conditioned stimulus arrives at the dendritic synapses, our learning mechanism would lead to ramping activity and salivation prior to the onset of the unconditioned stimulus that originally triggered the salivation. To our knowledge, there is no conclusive experimental data to support or discard the hypothesis that prospective coding is involved in classical trace conditioning. In the cited studies on ramping activity [1, 2, 6–11], the animals were actively engaged in a task (operant conditioning). It is unlikely, however, that the ramping activity is merely a side-effect of movement preparation, since Rainer et al. [1] found it to be stimulus-specific but not action-specific.

In our model of delayed paired-associate tasks, activity ramps rely on temporally structured input from short-term memory neurons. The usage of these short-term memory neurons is motivated by the observation that hippocampal activity is needed to overcome the temporal discontinuity in trace conditioning [4, 27]. We modeled the dynamics of the recurrent short-term memory network with a stochastic process. The parameter choice of this stochastic process is inspired by the widespread experimental observation that stimulus onset quenches the neural variability [28, 29]. It should also be possible to model the memory traces with “dynamical attractors” in recurrent networks of rate neurons [30] or with long and stable transient dynamics in balanced networks [31]. Since these memory traces are not the main focus of this study we generated them in a simpler way with the stochastic process, which still feels more natural than the delay-line like traces used in a study on trace conditioning [32].

In recurrent neural networks the learning rule of prospective coding allows fast replay of slow input sequences (Fig 6). Fast replay could be valuable for planning, where it is important to quickly assess the likely successors of a given state. The same fast replay of a previously induced slower activity sequence was also observed in the rat primary visual cortex [33] and it is as well studied as compressed hippocampal replay of a spatial sequence [34]. In rats these replay events can be observed minutes or hours after the spatial experience. In contrast, the simple form of the plasticity rule in Eq 8 does not have any consolidation properties and ongoing pre- and postsynaptic activity would quickly change the learned weight patterns and thus overwrite the memories. It is, however, straightforward to extend the plasticity model by a consolidation mechanism. In the three state consolidation model of Ziegler et al. [35, 36], early long-term potentiation (LTP) is induced by a triplet rule [37]. Replacing the triplet rule by the plasticity rule in Eq 8 would endow the learning rule of prospective coding with a consolidation mechanism. Such a consolidation mechanism would allow to replay sequences a long time after the training session.

Aiming at a better understanding of biological implementations of prediction learning, our model allows to speculate about physiological realizations of the model variables. Similar to previously proposed plasticity rules [16, 18], our learning mechanism depends on the postsynaptic firing rate $\varphi(U)$, a function of the dendritic potential $\varphi(V_w^*)$, the postsynaptic potential PSP and, as a new ingredient compared to previous propositions [16, 18]: a low-pass filtered version of the postsynaptic potential $\widetilde{\text{PSP}}$. A plasticity window that is slightly larger than the

duration of a postsynaptic potential is in agreement with experimentally measured plasticity window sizes [13, 38]. In particular, an increased level of dopamine was observed to expand the effective time window of potentiation to at least ~ 45 ms [38]. Importantly, even with a plasticity window on this timescale, predictions can be learned on a timescale of seconds due to the bootstrapping effect of predicting the own predictions.

We have shown that the proposed learning mechanism is closely related to temporal difference learning with eligibility traces $TD(\lambda)$. As discussed in the previous paragraph, a local biological implementation of our learning rule seems straightforward. In contrast, it seems more challenging to locally implement the delta error of TD learning. Potjans et al. and Kolodziejski et al. propose a local implementation that depends either on differential Hebbian plasticity [39] or on two postsynaptic activity traces with different time constants to approximate the difference in the delta error [40]. Both methods require a gating mechanism that allows plasticity only shortly after the onset of a new state and they require transition intervals between states of fixed duration. Furthermore, “state neurons” are only highly active when the agent is in a certain state, which requires the segmentation of the sensory input stream into discrete states. The learning rule we propose does not require these strong assumptions.

Frémaux et al. [41] speculate about a non-local implementation of TD learning with spiking neurons, where the TD error is represented by the firing rate of dopaminergic neurons that receive input from three groups of neurons that encode reward, value function and derivative of the value function. In the simulations, however, Frémaux et al. did not use the proposed network implementation of the TD error and they mention that it remains to be seen whether such a circuit can effectively be used to compute a useful TD error. A non-local implementation of the TD error appears compelling in a actor-critic setting, since the actor and the critic can be learned with the same TD signal. However, if the task is to predict more than a scalar quantity like reward, it seems inefficient to use a non-local implementation of the TD error for each quantity to be predicted. Already in our simple example of prospective coding in a recurrent neural network, four TD error networks would be needed in such a non-local implementation.

Generally, associating temporally separated events requires some memory of the first event until the second event is present. Possible neural implementations of this memory rely on long spiking activity traces or on long synaptic eligibility traces. Our model of the delayed paired-associate task relies on long spiking activity traces. The short-term memory network can be seen as a liquid state machine [42] or echo state machine [43] and the ramping activity is learned as readout from this activity traces. Alternatively, the activity trace could be represented by slowly, exponentially decaying spiking activity after strong stimulation of a cell [44]. This proposition, however, fails to explain the experimentally observed activity ramps prior to predictable events [1, 2, 6–11]

The origin of the ramping activity observed in experiments is not yet fully understood. An alternative to our proposition can be found in recurrent neural network dynamics, where slowly ramping or decaying activity arises with appropriately tuned synaptic weights [2, 25]. In a reinforcement learning setting the time constant of the ramp can be learned by adjusting the recurrent weights with reward modulated Hebbian plasticity [45]. Data analysis of recordings in the macaque lateral intraparietal area revealed yet another candidate explanation: single neuron activity profiles could follow a step-like time course, while the averaged activity is a ramp, if the steps occur at different points in time [46].

Despite the formal link of our prospective coding algorithm to TD learning, the learning we consider is purely supervised on the level of the neuron. Yet, the same learning rule can also be used to explain conditioning experiments. Instead of the multiplicative modulation by a global reward signal, the reward signal could directly nudge the somatic compartment of the

neurons and act as a teaching signal. But the learning rule would also allow for combining the somatic nudging signal with an additional modulatory factor, and nudging and modulatory signals could even be sparse and interleaved. For instance, the rule may explain the simultaneous shaping of predictive motor circuitries by sensory feedback and reward [5]. Fluctuating somatic inputs may cause behavioral variations and feedback signals may gate dendritic plasticity such that only rewarded fluctuations act as a target signal for prospective coding. It is also possible to adapt the somatic input connections directly with reinforcement learning, and a ramping activity could arise from learning a prospective code with stimulus-dependent dendritic input.

Since reward is an intrinsic component in animal training, we acquired an advanced knowledge about the neuronal bases of reward prediction. But predictions are not restricted to reward, and predicting the identity of stimuli yields more versatile information. We speculate that prospective coding is more abundant than previously thought and, as we showed, it could easily be implemented on the level of an individual neuron. This view is also consistent with the recently observed future-predicting encoding in the retina [47]. To this end, a potentiation window slightly larger than a PSP, together with the bootstrapping effect of predicting the own predictions, is a parsimonious mechanism for learning prospective codes by neurons. A characteristic of these neurons is that their current firing rate matches their own expected future discounted rate.

Methods

Parameters of the neuron model

The spike response kernel κ in Eq 2 is given by

$$\kappa(t) = cH(t)(e^{-t/\tau_m} - e^{-t/\tau_s}), \tag{20}$$

with Heaviside function $H(t) = 0$ if $t < 0$ and $H(t) = 1$ otherwise, $\tau_m = 10$ ms, $\tau_s = 10/3$ ms and $c^{-1} = \int_{-\infty}^{\infty} dt H(t)(e^{-t/\tau_m} - e^{-t/\tau_s})$.

We set the somatic capacitance $C = 1$ nF, the leak conductance $g_L = 100$ nS, the coupling conductance $g_D = 1.8 \mu S$, and the excitatory and inhibitory reversal potential $E_E = 14/3$ and $E_I = -1/3$, respectively. The description of the excitatory conductance $g_E(t)$ is given in the figure captions. The inhibitory nudging conductance $g_I(t)$ was equal to 0 except for simulations with $\widetilde{PSP} = PSP$ in Fig 2, where $g_I(t) = 4g_E(t)$. The resting potential is 0 for both, the dendritic potential V_w and the somatic potential U . If one takes our unitless resting potential of 0 to correspond to -70 mV, and a potential 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).

The instantaneous firing rate of the neuron is assumed to depend on the somatic membrane potential through function $\varphi(U)$, which in the simulations has the form

$$\varphi(U) = \begin{cases} \varphi_{\max} & \text{if } U > 1 \\ 0 & \text{if } U < 0 \\ \varphi_{\max} \cdot U & \text{otherwise,} \end{cases} \tag{21}$$

with $\varphi_{\max} = 0.06$ kHz. In simulations with spiking, the firing rate multiplied by the simulation time step serves as instantaneous rate of an inhomogeneous Bernoulli process.

Steady state solution of the somatic potential dynamics

For slowly enough changing $I_{\text{tot}}(t)$ and $g_{\text{tot}}(t)$, $U(t)$ is well approximated by $I_{\text{tot}}(t)/g_{\text{tot}}(t)$. To see this, we use the ansatz $U(t) = I_{\text{tot}}(t)/g_{\text{tot}}(t) + \epsilon(t)$ in Eq 1 and find

$$\dot{\epsilon} = -\frac{g_{\text{tot}}}{C}\epsilon - \frac{\dot{I}_{\text{tot}}}{g_{\text{tot}}} + \frac{\dot{g}_{\text{tot}}I_{\text{tot}}}{g_{\text{tot}}^2}, \tag{22}$$

which leads to the conclusion that the error ϵ is small if $\left| \frac{\dot{I}_{\text{tot}}}{g_{\text{tot}}} + \frac{\dot{g}_{\text{tot}}I_{\text{tot}}}{g_{\text{tot}}^2} \right| \ll 1$ during at least an interval of approximate duration $\frac{C}{g_{\text{tot}}}$.

Under these assumptions we write

$$U(t) \approx \frac{I_{\text{tot}}(t)}{g_{\text{tot}}(t)} = \lambda(t)V_w^*(t) + U^*(t), \tag{23}$$

where we introduce the ‘nudging’ factor $\lambda(t) = \frac{g_L + g_D}{g_{\text{tot}}(t)}$, the attenuated dendritic potential

$$V_w^*(t) = \frac{g_D}{g_L + g_D} V_w(t), \text{ and the attenuated somatic input } U^*(t) = \frac{g_E(t)E_E + g_I(t)E_I}{g_{\text{tot}}(t)}.$$

Generality of the results

Our main results are robust to variations of the model. For example, the target input I could be given by the input through static synapses on another dendritic branch instead of synapses at the soma, i.e. $I(t) = g_D(V_s(t) - U)$. In this case, the nudging factor becomes $\lambda = \frac{g_L + g_D}{g_L + 2g_D}$ and is constant in time.

Modifying the depression term of the learning rule has an effect on the effective time scale τ_{eff} , but large effective time constants are achievable in any case. If the depression term in Eq 8 would be replaced by $-\varphi(\lambda V_w^*)\text{PSP}_i$, the effective time constant would be $\tau_{\text{eff}} \approx \tau/(1 - \alpha)$, i.e. τ_{eff} would be independent of λ but still diverge when $\alpha \rightarrow 1$. Similarly, for a depression term given by $-\varphi(V_w)\text{PSP}_i$, the effective time constant would be $\tau_{\text{eff}} \approx \tau/(1 - \lambda_2 \alpha)$, with $\lambda_2 = g_D/g_{\text{tot}}$.

In the current writing of the learning rule, Eq 8, the postsynaptic term arises as instantaneous firing rate $\varphi(U)$. But this rate could also be replaced by a postsynaptic sample spike train $S(t)$ that averages out to this same rate, $\langle S(t) \rangle = \varphi(U(t))$. Since learning becomes slower by this sampling, we run our simulations in the form of Eq 8.

Dynamics of short-term memory neurons in the delayed paired associate task

For each STM neuron i we first choose template rate trajectories $r_i^1(t)$ for stimulus A1 and $r_i^2(t)$ for stimulus A2 by sampling from a mean-zero Ornstein-Uhlenbeck process

$$dr_i^s(t) = -\theta_1 r_i^s(t)dt + \sigma_1 dW(t), \tag{24}$$

where W is a Wiener process, $1/\theta_1 = 1000$ ms, $\sigma_1 = 1$ and $s \in \{1, 2\}$. Actual rate trajectories $r_i(t)$ were sampled from a process with trial dependent mean and time dependent variance, i.e.

$$dr_i(t) = -\theta_2 (r_i(t) - \mu^s(t))dt + \sigma(t)dW(t), \tag{25}$$

where $1/\theta_2 = 100$ ms,

$$\mu^s(t) = (1 - \sigma(t))r_i^s(t) - \sigma(t)/2 \tag{26}$$

and $\sigma^2(t) = 1$ if $t < 0$ s or $t > 3$ s, $\sigma^2(t) = 0.1$ otherwise. This assures that in each trial the rate trajectories approach the template trajectories during the presentation of the sample. In between

trials, the rate trajectories are independent of the template trajectories. Spike times are determined by sampling from an inhomogeneous Bernoulli process with rate $\varphi(r_i(t))\Delta t$, where Δt is the simulation time step.

Simulation details

The differential equations were integrated with the Euler forward method with step size 0.1 ms. We choose the learning rate $\eta = 0.5$ in all simulations except for the simulation in Fig 2B and 2C where $\eta = 50$, since the presynaptic firing rate is low. All simulations are written in C. The plots are generated with Mathematica. The source code is publicly available at <https://github.com/jbrea/prospectiveCoding>.

Stationary point of learning for periodic environments

We assume a stationary environment and rich dendritic input dynamics, such that the dendritic inputs can potentially be predictive of the somatic input. There are different ways to model stationarity of the environment. One way is to restrict the inputs to depend on a stationary latent Markov chain. We consider this case in detail in the next section. Here, to present the main ideas in a mathematically simple form, we look at the artificial case, where stationarity enters through deterministic and periodic functions $\text{PSP}_i(t)$ and $U^*(t)$ with period T . Under this assumption, learning is at a stationary point when the changes of the weights in Eq 8 integrated over one period vanish, i.e.

$$0 = \int_0^T dt \dot{w}_i(t) = \eta \int_0^T dt \left(\alpha \varphi(U(t)) \widetilde{\text{PSP}}_i(t) - \varphi(V_w^*(t)) \text{PSP}_i(t) \right). \quad (27)$$

Using the definition of $\widetilde{\text{PSP}}$ in Eq 9 we find

$$0 = \int_0^T dt \left(\alpha \varphi(U(t)) \frac{1}{\tau} \int_0^\infty ds e^{-\frac{s}{\tau}} \text{PSP}_i(t-s) - \varphi(V_w^*(t)) \text{PSP}_i(t) \right) \quad (28)$$

$$= \int_0^T dt \left(\frac{\alpha}{\tau} \int_0^\infty ds e^{-\frac{s}{\tau}} \varphi(U(t+s)) - \varphi(V_w^*(t)) \right) \text{PSP}_i(t), \quad (29)$$

where Eq 29 is obtained by changing the order of integration, changing the integration variable t to $t+s$ and using $\int_{-s}^{T-s} dt f(t) = \int_0^T dt f(t)$, which holds for any T -periodic function $f(t)$. The puzzling transition from an integral that depends on the past values of PSP_i in Eq 28 to an integral that depends on the future values of U in Eq 29 is a result of the assumed stationarity of the environment, which here is expressed in the periodicity of the functions $\text{PSP}_i(t)$ and $U^*(t)$. Eq 29 holds for all synapses i , if

$$\varphi(V_w^*(t)) = \frac{\alpha}{\tau} \int_0^\infty ds e^{-\frac{s}{\tau}} \varphi(U(t+s)). \quad (30)$$

Strictly, Eq 30 follows from Eq 29 only if the inputs $\text{PSP}_i(t)$ span the space of square integrable, T -periodic functions. In actual implementations the number of synapses is limited, but we find empirically that Eq 30 holds approximately at the stationary point if, loosely speaking, the inputs $\text{PSP}_i(t)$ at individual synapses are sufficiently rich and different from each other.

The right-hand side of Eq 30 also depends on the dendritic potential V_w^* , since the membrane potential U depends both on the dendritic input V_w^* and the somatic input U^* (see Eq 4).

Assuming a linear transfer function φ , Eq 30 becomes

$$\varphi(V_w^*(t)) = \frac{\alpha}{\tau} \int_0^\infty ds e^{-\frac{s}{\tau}} (\lambda \varphi(V_w^*(t+s)) + \varphi(U^*(t+s))). \quad (31)$$

With a Fourier transform and assuming a constant nudging factor λ we can solve this equation for $\varphi(V_w^*(t))$.

The Fourier coefficients \hat{h}_k , $k \in \mathbb{Z}$, of the T -periodic function $h(t) = \int_0^\infty ds e^{-\frac{s}{\tau}} f(t+s)$ are given by

$$\hat{h}_k = \int_0^T dt e^{-i\frac{2\pi k t}{T}} \int_0^\infty ds e^{-\frac{s}{\tau}} f(t+s) = \int_0^T dt e^{-i\frac{2\pi k t}{T}} f(t) \int_0^\infty ds e^{\frac{2\pi k s}{T}} e^{-\frac{s}{\tau}} \quad (32)$$

$$= \hat{f}_k \int_0^\infty ds e^{s(\frac{2\pi k}{T} - \frac{1}{\tau})} \quad (33)$$

$$= \hat{f}_k \frac{1}{\frac{1}{\tau} - i\frac{2\pi k}{T}}, \quad (34)$$

where, in the first line, we changed the order of integration, changed the variable t to $t-s$ and used the periodicity of the integrand to obtain $\int_{-s}^{T-s} dt e^{-i\frac{2\pi k t}{T}} f(t) = \int_0^T dt e^{-i\frac{2\pi k t}{T}} f(t)$. In the second line we introduced the Fourier coefficients \hat{f}_k .

With $f(t) = \varphi(V_w^*(t))$ and $g(t) = \varphi(U^*(t))$ we rewrite Eq 31

$$f(t) = \frac{\alpha}{\tau} \int_0^\infty ds e^{-\frac{s}{\tau}} (\lambda f(t+s) + g(t+s)) \quad (35)$$

and Fourier transform both sides to obtain

$$\hat{f}_k = \frac{\alpha}{\tau} (\lambda \hat{f}_k + \hat{g}_k) \frac{1}{\frac{1}{\tau} - i\frac{2\pi k}{T}}. \quad (36)$$

Solving for \hat{f}_k leads to

$$\hat{f}_k = \frac{\alpha}{\tau} \hat{g}_k \frac{\frac{1}{\frac{1}{\tau} - i\frac{2\pi k}{T}}}{1 - \frac{\alpha\lambda}{\tau} \frac{1}{\frac{1}{\tau} - i\frac{2\pi k}{T}}} \quad (37)$$

$$= \frac{\alpha}{\tau} \hat{g}_k \frac{1}{\frac{1-\alpha\lambda}{\tau} - i\frac{2\pi k}{T}}. \quad (38)$$

This equation has the same structure as Eq 34. With the inverse Fourier transform and assuming $\alpha\lambda < 1$ we find Eq 10, i.e.

$$f(t) = \frac{\alpha}{\tau} \int_0^\infty ds e^{-\frac{s}{\tau_{\text{eff}}}} g(t+s), \quad (39)$$

where $\tau_{\text{eff}} = \frac{\tau}{1-\alpha\lambda}$.

Convergence of learning in stationary stochastic environments

We formalize the notion of a stationary environment by introducing a stationary latent Markov chain and restricting the dendritic input $\text{PSP}_i(t) = \text{PSP}_i(X_t)$ and the somatic input $U^*(t) = U^*(X_t)$ to depend on the state X_t of the Markov chain. An alternative way to formalize the notion of

stationarity would be to define stationary dynamics of the dendritic inputs and define the correlation between dendritic and somatic input. As it is always possible to reformulate the stationary dendritic input dynamics and the correlation between dendritic and somatic input in terms of a stationary latent Markov chain—with potentially large state space—we stick to the description with a latent Markov chain.

Formally, for time $t \in \mathbb{Z}$, states X_t in a finite set $\mathcal{X} = (s_1, s_2, \dots, s_N)$ evolve according to a stationary, irreducible Markov chain with transition probabilities $T(s_i, s_j) = Pr(X_{t+1} = s_j | X_t = s_i)$ and stationary distribution $\pi(s_i) = Pr(X_t = s_i)$.

Note that the case of deterministic periodic input is readily formulated in terms of a stationary latent Markov chain that cycles deterministically through the state space, e.g. $T(s_i, s_j) = 1$ if $j = i + 1$ or $j = 1$ and $i = N$ and $T(s_i, s_j) = 0$ otherwise. Functions that depend only on the state of the Markov chain are thus cyclic with period N , e.g. $PSP_i(X_t) = PSP_i(X_{t+N})$.

In order to switch to matrix notation in the rest of this section, we introduce the following terms:

- Discounting operator $A = \alpha \sum_{t=0}^{\infty} \gamma^t T^t$, with transition matrix T and discount factor $\gamma \in [0, 1)$.
- Postsynaptic potentials $\mathbf{b}_i = (PSP_i(s_1), \dots, PSP_i(s_N))'$ for each synapse i .
- Matrix of postsynaptic potentials $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_S]$, where S is the number of synapses.
- Postsynaptic firing rates $\mathbf{r}_U = (\varphi(U(s_1)), \dots, \varphi(U(s_N)))'$.
- Dendritic rates $\mathbf{r}_V = (\varphi(V_w^*(s_1)), \dots, \varphi(V_w^*(s_N)))'$.
- Somatic input rates $\mathbf{r}_I = (\varphi(U^*(s_1)), \dots, \varphi(U^*(s_N)))'$.
- Expected future discounted firing rate $F(x) = (A\mathbf{r}_U)(x) = \alpha \sum_{t=0}^{\infty} \gamma^t \mathbb{E}[\varphi(U(X_t)) | X_0 = x]$.
- Expected low-pass filtered postsynaptic potential $\widetilde{PSP}_i(x) = \sum_{t=0}^{\infty} \gamma^t \mathbb{E}[PSP_i(X_{-t}) | X_0 = x]$.

In the following we sketch the proof for the equivalents of Eqs 30, 12 and 11 in the Markov chain setting. We will make use of the following basic facts about conditional expectations:

$$\mathbb{E}[f(X_t) | X_0 = x] = \sum_{x_0, x_1, \dots, x_t} \delta(x, x_0) \prod_{s=1}^t T(x_{s-1}, x_s) f(x_t) = (T^t \mathbf{f})(x), \tag{40}$$

$$\mathbb{E}[f(X_{-t}) | X_0 = x] = \frac{1}{\pi(x)} \sum_{x_{-t}} f(x_{-t}) \pi(x_{-t}) Pr(X_0 = x | X_{-t} = x_{-t}) = \frac{1}{\pi(x)} (\mathbf{f}' \Pi T^t)(x), \tag{41}$$

where $t > 0$, T^t denotes the matrix power of T , $\Pi = \text{diag}(\pi(s_1), \pi(s_2), \dots, \pi(s_N))$ is the diagonal “stationary distribution matrix”, column vector $\mathbf{f} = (f(s_1), f(s_2), \dots, f(s_N))'$ and row vector \mathbf{f}' , the transposed of \mathbf{f} .

1. At the fixed point of learning the dendritic rate is proportional to the expected future discounted firing rate (cf. Eq 30). In discrete time the learning rule in Eq 8 becomes

$$\Delta w_{t,i} = \eta \left(\alpha \varphi(U(X_t)) \widetilde{PSP}_{t,i} - \varphi(V_w^*(X_t)) PSP_i(X_t) \right), \tag{42}$$

with $\widetilde{PSP}_{t,i} = \sum_{s=0}^{\infty} \gamma^s PSP_i(X_{t-s})$. While $\Delta w_{t,i}$ is a stochastic variable in general, we will discuss in the following only the corresponding ordinary differential equation (ODE) of the mean

$$\dot{w}_i = \eta \sum_{x \in \mathcal{X}} \pi(x) \left(\alpha \varphi(U(x)) \widetilde{PSP}_i(x) - \varphi(V_w^*(x)) PSP_i(x) \right), \tag{43}$$

where $\widetilde{\text{PSP}}_i(x) = \sum_{t=0}^{\infty} \gamma^t \mathbb{E}[\text{PSP}(X_{-t}) | X_0 = x]$ is the expected low-pass filtered postsynaptic potential. This ODE has the same fixed point and convergence behavior as the dynamics in Eq 42 under mild assumptions [48].

As in Eqs 28 and 29, we are going to show now that we can rewrite the dynamics of the mean synaptic weight in terms of the future discounted firing rate $F(x)$ instead of the expected low-pass filtered postsynaptic potential $\widetilde{\text{PSP}}_i(x)$, i.e.

$$\dot{w}_i = \eta \sum_x \pi(x) (F(x) - \varphi(V_w^*(x))) \text{PSP}_i(x). \tag{44}$$

This result is a consequence of the assumed stationarity of the Markov chain. It can be found by focusing on the potentiation term in the learning rule in Eq 43 and using Eq 41 and the notation introduced in the last paragraph, in particular,

$$\widetilde{\text{PSP}}_i(x) = \sum_{t=0}^{\infty} \gamma^t \mathbb{E}[\text{PSP}(X_{-t}) | X_0 = x] = \frac{1}{\pi(x)} \left(\sum_{t=0}^{\infty} \gamma^t \mathbf{b}' \Pi T^t \right) (x) = \frac{1}{\pi(x)\alpha} (\mathbf{b}' \Pi A)(x), \tag{45}$$

which leads to

$$\sum_x \pi(x) \alpha \varphi(U(x)) \widetilde{\text{PSP}}_i(x) = \mathbf{b}'_i \Pi A \mathbf{r}_U = \sum_x \pi(x) (A \mathbf{r}_U)(x) \text{PSP}_i(x) = \sum_x \pi(x) F(x) \text{PSP}_i(x). \tag{46}$$

Using this equality in Eq 43 leads to Eq 44.

Assuming a trivial kernel for $B\Pi$, i.e. $B\Pi \mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{x} = \mathbf{0}$, we find by looking at Eq 44 that

$$\forall i : \dot{w}_i = 0 \Leftrightarrow \forall x : \varphi(V_w^*(x)) = F(x), \tag{47}$$

which is analogous to the statement in Eq 30. The assumption of a trivial kernel of $B\Pi$ implies that the map $\text{PSP}(s_i)$ from the state space of the latent Markov chain to dendritic inputs is one-to-one. This assumption is analogous to the statement that the dendritic inputs $\text{PSP}_i(t)$ at individual synapses are sufficiently rich and different from each other (see Results after Eq 30).

2. At the fixed point of learning the dendritic rate that is proportional to the expected future discounted somatic input rate, but with a longer discount time constant (cf. Eq 12). Since the future discounted firing rate $F(x)$ depends on $\varphi(V_w^*(x))$, it is not trivial to solve Eq 47 for $\varphi(V_w^*(x))$. Similar as in the result section in Eqs 31 and 10, however, we can show for a linear φ and constant $\lambda(s_i) = \lambda$ that

$$\varphi(V_w^*(x)) = \frac{\alpha}{1 - \lambda\alpha} \sum_{t=0}^{\infty} \gamma_{\text{eff}}^t \mathbb{E}[\varphi(U^*(X_t)) | X_0 = x], \tag{48}$$

with $\gamma_{\text{eff}} = \frac{\gamma}{1 - \lambda\alpha}$. Indeed, assuming linear φ we can rewrite Eq 47 in vector notation and solve for \mathbf{r}_V to obtain

$$\mathbf{r}_V = A \mathbf{r}_U = A(\lambda \mathbf{r}_V + \mathbf{r}_I) \tag{49}$$

$$\Rightarrow \mathbf{r}_V = (1 - \lambda A)^{-1} A \mathbf{r}_I = \sum_{s=0}^{\infty} \lambda^s A^{s+1} \mathbf{r}_I, \tag{50}$$

where we have assumed that $\lambda\alpha < 1 - \gamma$ such that the series converges. Powers of A evaluate to

$$A^{s+1} = \alpha^{s+1} \sum_{t_1=0, \dots, t_{s+1}=0}^{\infty} \gamma^{t_1 + \dots + t_{s+1}} T^{t_1 + \dots + t_{s+1}} = \alpha^{s+1} \sum_{t=0}^{\infty} \binom{t+s}{t} \gamma^t T^t, \tag{51}$$

and thus we can rewrite Eq 50 to get,

$$\begin{aligned} \mathbf{r}_V &= \alpha \sum_{t=0}^{\infty} \underbrace{\sum_{s=0}^{\infty} (\lambda\alpha)^s \binom{t+s}{t}}_{=(1-\lambda\alpha)^{-t-1}} \gamma^t T^t \mathbf{r}_I \\ &= \frac{\alpha}{1-\lambda\alpha} \sum_{t=0}^{\infty} \left(\frac{\gamma}{1-\lambda\alpha}\right)^t T^t \mathbf{r}_I, \end{aligned} \tag{52}$$

which proves the claim in Eq 48. The effective time constant $\gamma_{\text{eff}} = \frac{\gamma}{1-\lambda\alpha}$ can be much larger than γ . In fact, for $\alpha \rightarrow (1-\gamma)/\lambda$ we find $\gamma_{\text{eff}} \rightarrow 1$.

Remarks.

1. For affine $\varphi(u) = a \cdot u + c$ the equivalent of Eq 50 is

$$\mathbf{r}_V = (1 - \lambda A)^{-1} A(\mathbf{r}_I - \mathbf{c}). \tag{53}$$

In a first order approximation, the stationary \mathbf{r}_V for a non-linear φ is thus a translated version of the stationary solution for a linear φ .

2. For input $\mathbf{r}_I = 0$ and linear φ we find that at the stationary point of learning $\mathbf{r}_V = 0$. Thus we expect that learned weights decay again, once input \mathbf{r}_I is removed.

3. Convergence of learning. For linear $\varphi(u) = \phi \cdot u$ with $\phi > 0$, we have $\mathbf{r}_V = \phi B \mathbf{w}$ and thus the learning rule in Eq 43 can be written in vector notation as

$$\dot{\mathbf{w}} = \eta B' \Pi (A(\Lambda \mathbf{r}_V + \mathbf{r}_I) - \mathbf{r}_V) = \underbrace{\eta \phi B' \Pi (A\Lambda - 1) B \mathbf{w}}_{=X} + \underbrace{\eta B' \Pi A \mathbf{r}_I}_{=\mathbf{c}}, \tag{54}$$

where we introduced the diagonal ‘nudging matrix’ $\Lambda = \text{diag}(\lambda(s_1), \dots, \lambda(s_N))$.

With $\mathbf{w}^* = -X^+ \mathbf{c} + (1 - X^+ X) \mathbf{w}$ the orthogonal projection of \mathbf{w} onto $W^* = \{\mathbf{w} | X \mathbf{w} = -\mathbf{c}\}$, where X^+ denotes the Moore-Penrose pseudoinverse of X , we are going to show that $L(\mathbf{w}) = \frac{1}{2} (\mathbf{w} - \mathbf{w}^*)^2$ is a Lyapunov function of the dynamics in Eq 54. With $\mathbf{y} = \mathbf{w} - \mathbf{w}^*$ and $\mathbf{z} = \sqrt{\eta \phi} B \mathbf{y}$, the temporal evolution of L is given by

$$\begin{aligned} \dot{L} &= (\nabla L)' \dot{\mathbf{w}} \\ &= \mathbf{y}' (X(\mathbf{w}^* + \mathbf{y}) + \mathbf{c}) \\ &= \mathbf{y}' X \mathbf{y} \\ &= \eta \phi \mathbf{y}' B' \Pi (A\Lambda - 1) B \mathbf{y} \\ &= \langle \mathbf{z}, A\Lambda \mathbf{z} \rangle - \langle \mathbf{z}, \mathbf{z} \rangle \leq 0, \end{aligned}$$

where we defined the scalar product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}' \Pi \mathbf{y}$ and the inequality follows since both A and Λ are contracting maps, i.e. $\sqrt{\langle A\mathbf{z}, A\mathbf{z} \rangle} = \|A\mathbf{z}\| \leq \|\mathbf{z}\|$ and therefore

$$\langle \mathbf{z}, A\Lambda \mathbf{z} \rangle \leq \|\mathbf{z}\| \|A\Lambda \mathbf{z}\| \leq \|\mathbf{z}\| \|\Lambda \mathbf{z}\| \leq \|\mathbf{z}\| \|\mathbf{z}\| = \langle \mathbf{z}, \mathbf{z} \rangle. \tag{55}$$

Λ is contracting because it is diagonal with entries between 0 and 1 and A is contracting

because

$$\|Az\|^2 = \sum_i \pi(i) \left(\sum_j A_{ij} z_j \right)^2 \leq \sum_i \pi(i) \underbrace{\left(\sum_j A_{ij}^2 \right)}_{<1} \left(\sum_j z_j^2 \right) \leq \sum_i \pi(i) \sum_j z_j^2 = \|z\|^2, \quad (56)$$

where we used the facts that $0 \leq A_{ij} < 1$ and the row sums of A are equal to $\frac{\gamma}{1-\gamma}$ and therefore $\sum_j A_{ij}^2 \leq \sum_j A_{ij} < 1$.

Relation between the learning rule in Eq 42 and TD(1)

For $\lambda_{TD} = 1$ and therefore $\widetilde{PSP} = \widehat{PSP}$, we can rewrite Eq 17 by expanding the delta error in Eq 19 and using the identity $\gamma \widehat{PSP}_{t,i} = \widehat{PSP}_{t+1,i} - PSP_i(X_{t+1})$ to find

$$\begin{aligned} \delta_t \widehat{PSP}_{t,i} &= \alpha \varphi(U(X_t)) \widehat{PSP}_{t,i} + \varphi(V_{w_t}^*(X_{t+1})) (\widehat{PSP}_{t+1,i} - PSP_i(X_{t+1})) - \varphi(V_{w_t}^*(X_t)) \widehat{PSP}_{t,i} \\ &= \alpha \varphi(U(X_t)) \widehat{PSP}_{t,i} - \varphi(V_{w_t}^*(X_{t+1})) PSP_i(X_{t+1}) \end{aligned} \quad (57)$$

$$+ \varphi(V_{w_t}^*(X_{t+1})) \widehat{PSP}_{t+1,i} - \varphi(V_{w_t}^*(X_t)) \widehat{PSP}_{t,i}. \quad (58)$$

With small parameter updates in each time step, the terms in Eq 58 approximately cancel each other when summing over subsequent terms: $\delta_t \widehat{PSP}_{t,i}$ contributes $+\varphi(V_{w_t}^*(X_{t+1})) \widehat{PSP}_{t+1,i}$ and $\delta_{t+1} \widehat{PSP}_{t+1,i}$ contributes $-\varphi(V_{w_{t+1}}^*(X_{t+1})) \widehat{PSP}_{t+1,i}$. What remains are the terms in Eq 57, which resemble the terms in the learning rule in Eq 42.

Acknowledgments

We thank Christian Pozzorini, Laureline Logiaco and Kristin Völk for valuable comments on the manuscript.

Author Contributions

Conceived and designed the experiments: RU WS JB. Performed the experiments: JB. Analyzed the data: JB WS. Wrote the paper: JB WS ATG. Methods: JB ATG RU WS.

References

1. Rainer G, Rao SC, Miller EK (1999) Prospective Coding for Objects in Primate Prefrontal Cortex. *The Journal of Neuroscience* 19: 5493–5505. PMID: [10377358](#)
2. Reutimann J, Yakovlev V, Fusi S, Senn W (2004) Climbing Neuronal Activity as an Event-Based Cortical Representation of Time. *Journal of Neuroscience* 24: 3295–3303. doi: [10.1523/JNEUROSCI.4098-03.2004](#) PMID: [15056709](#)
3. Pavlov IP (1927). *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. Oxford Univ. Press
4. Bangasser DA, Waxler DE, Santollo J, Shors TJ (2006) Trace conditioning and the hippocampus: the importance of contiguity. *The Journal of Neuroscience* 26: 8702–6. doi: [10.1523/JNEUROSCI.1742-06.2006](#) PMID: [16928858](#)
5. Li N, Chen Tw, Guo ZV, Gerfen CR, Svoboda K (2016) A motor cortex circuit for motor planning and movement. *Nature* 519: 51–56. doi: [10.1038/nature14178](#)
6. Liu Z, Zhou J, Li Y, Hu F, Lu Y, et al. (2014) Dorsal raphe neurons signal reward through 5-HT and glutamate. *Neuron* 81: 1360–74. doi: [10.1016/j.neuron.2014.02.010](#) PMID: [24656254](#)
7. Howe MW, Tierney PL, Sandberg SG, Phillips PEM, Graybiel AM (2013) Prolonged dopamine signaling in striatum signals proximity and value of distant rewards. *Nature* 500: 575–9. doi: [10.1038/nature12475](#) PMID: [23913271](#)

8. van der Meer MAA, Redish AD (2011) Theta phase precession in rat ventral striatum links place and reward information. *The Journal of Neuroscience* 31: 2843–2854. doi: [10.1523/JNEUROSCI.4869-10.2011](https://doi.org/10.1523/JNEUROSCI.4869-10.2011) PMID: [21414906](https://pubmed.ncbi.nlm.nih.gov/21414906/)
9. Quintana J, Fuster JM (1999) From perception to action: Temporal integrative functions of prefrontal and parietal neurons. *Cerebral Cortex* 9: 213–221. doi: [10.1093/cercor/9.3.213](https://doi.org/10.1093/cercor/9.3.213) PMID: [10355901](https://pubmed.ncbi.nlm.nih.gov/10355901/)
10. Miller EK, Erickson Ca, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *The Journal of Neuroscience* 16: 5154–5167. PMID: [8756444](https://pubmed.ncbi.nlm.nih.gov/8756444/)
11. Sakai K, Miyashita Y (1991) Neural organization for the long-term memory of paired associates. *Nature* 354: 152–155. doi: [10.1038/354152a0](https://doi.org/10.1038/354152a0) PMID: [1944594](https://pubmed.ncbi.nlm.nih.gov/1944594/)
12. Markram H, Lübke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275: 213–5. doi: [10.1126/science.275.5297.213](https://doi.org/10.1126/science.275.5297.213) PMID: [8985014](https://pubmed.ncbi.nlm.nih.gov/8985014/)
13. Bi GQ, Poo M (2001) Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu Rev Neurosci* 24: 139–66. doi: [10.1146/annurev.neuro.24.1.139](https://doi.org/10.1146/annurev.neuro.24.1.139) PMID: [11283308](https://pubmed.ncbi.nlm.nih.gov/11283308/)
14. Cichon J, Gan W (2015) Branch-specific dendritic Ca²⁺ spikes cause persistent synaptic plasticity. *Nature* 520: 180–5 doi: [10.1038/nature14251](https://doi.org/10.1038/nature14251) PMID: [25822789](https://pubmed.ncbi.nlm.nih.gov/25822789/)
15. Yagishita S, Hayashi-Takagi A, Ellis-Davies GC, Urakubo H, Ishii S, et al. (2014) A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science* 345: 1616–1620. doi: [10.1126/science.1255514](https://doi.org/10.1126/science.1255514) PMID: [25258080](https://pubmed.ncbi.nlm.nih.gov/25258080/)
16. Pfister JP, Toyozumi T, Barber D, Gerstner W (2006) Optimal spike-timing-dependent plasticity for precise action potential firing in supervised learning. *Neural Computation* 18: 1318–1348. doi: [10.1162/neco.2006.18.6.1318](https://doi.org/10.1162/neco.2006.18.6.1318) PMID: [16764506](https://pubmed.ncbi.nlm.nih.gov/16764506/)
17. Brea J, Senn W, Pfister JP (2013) Matching recall and storage in sequence learning with spiking neural networks. *The Journal of Neuroscience* 33: 9565–75. doi: [10.1523/JNEUROSCI.4098-12.2013](https://doi.org/10.1523/JNEUROSCI.4098-12.2013) PMID: [23739954](https://pubmed.ncbi.nlm.nih.gov/23739954/)
18. Urbanczik R, Senn W (2014) Learning by the Dendritic Prediction of Somatic Spiking. *Neuron* 81: 521–528. doi: [10.1016/j.neuron.2013.11.030](https://doi.org/10.1016/j.neuron.2013.11.030) PMID: [24507189](https://pubmed.ncbi.nlm.nih.gov/24507189/)
19. Wise S, Boussaoud D, Johnson P, Caminiti R (1997) Premotor and parietal cortex: corticocortical connectivity and combinatorial computations. *Annu Rev Neurosci* 20: 25–42. doi: [10.1146/annurev.neuro.20.1.25](https://doi.org/10.1146/annurev.neuro.20.1.25) PMID: [9056706](https://pubmed.ncbi.nlm.nih.gov/9056706/)
20. Jaakkola T, Jordan MI, Singh S (1994) On the Convergence of Stochastic Iterative Dynamic Programming Algorithms. *Neural Computation* 6: 1185–1201. doi: [10.1162/neco.1994.6.6.1185](https://doi.org/10.1162/neco.1994.6.6.1185)
21. Dayan P (1992) The convergence of TD(λ) for general λ . *Machine Learning* 8: 341–362. doi: [10.1007/BF00992701](https://doi.org/10.1007/BF00992701)
22. Sutton RS (1988) Learning to Predict by the Methods of Temporal Differences. *Machine Learning* 3: 9–44. doi: [10.1023/A:1022633531479](https://doi.org/10.1023/A:1022633531479)
23. Urbanczik R, Senn W (2009) Reinforcement learning in populations of spiking neurons. *Nature Neuroscience* 12: 250–2 doi: [10.1038/nn.2264](https://doi.org/10.1038/nn.2264) PMID: [19219040](https://pubmed.ncbi.nlm.nih.gov/19219040/)
24. Friedrich J, Urbanczik R, Senn W (2011) Spatio-temporal credit assignment in neuronal population learning. *PLoS Computational Biology* 7: e1002092. doi: [10.1371/journal.pcbi.1002092](https://doi.org/10.1371/journal.pcbi.1002092) PMID: [21738460](https://pubmed.ncbi.nlm.nih.gov/21738460/)
25. Gavornik J, Shuler M, Loewenstein Y, Bear M, Shouval HZ (2009) Learning reward timing in cortex through reward dependent expression of synaptic plasticity. Supporting Material. *Proceedings of the National Academy of Sciences* 106: 6826. doi: [10.1073/pnas.0901835106](https://doi.org/10.1073/pnas.0901835106)
26. He K, Huertas M, Hong S, Tie X, Hell J, et al. (2015) Distinct Eligibility Traces for LTP and LTD in Cortical Synapses. *Neuron*. doi: [10.1016/j.neuron.2015.09.037](https://doi.org/10.1016/j.neuron.2015.09.037)
27. Solomon PR, Vander Schaaf ER, Thompson RF, Weisz DJ (1986) Hippocampus and trace conditioning of the rabbit's classically conditioned nictitating membrane response. *Behavioral Neuroscience* 100: 729–744. doi: [10.1037/0735-7044.100.5.729](https://doi.org/10.1037/0735-7044.100.5.729) PMID: [3778636](https://pubmed.ncbi.nlm.nih.gov/3778636/)
28. Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, et al. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature neuroscience* 13: 369–78. doi: [10.1038/nn.2501](https://doi.org/10.1038/nn.2501) PMID: [20173745](https://pubmed.ncbi.nlm.nih.gov/20173745/)
29. Churchland MM, Abbott LF (2012) Two layers of neural variability. *Nature Neuroscience* 15: 1472–1474. doi: [10.1038/nn.3247](https://doi.org/10.1038/nn.3247) PMID: [23103992](https://pubmed.ncbi.nlm.nih.gov/23103992/)
30. Laje R, Buonomano DV (2013) Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nature Neuroscience* 16: 925–933. doi: [10.1038/nn.3405](https://doi.org/10.1038/nn.3405) PMID: [23708144](https://pubmed.ncbi.nlm.nih.gov/23708144/)

31. Hennequin G, Vogels T, Gerstner W (2014) Optimal Control of Transient Dynamics in Balanced Networks Supports Generation of Complex Movements. *Neuron* 82: 1394–1406. doi: [10.1016/j.neuron.2014.04.045](https://doi.org/10.1016/j.neuron.2014.04.045) PMID: [24945778](https://pubmed.ncbi.nlm.nih.gov/24945778/)
32. Ludvig EEA, Sutton RRS, Verbeek E, Kehoe EJ (2009) A computational model of hippocampal function in trace conditioning. In: *Advances in Neural Information Processing Systems 21*, Curran Associates, Inc. pp. 993–1000.
33. Xu S, Jiang W, Poo MM, Dan Y (2012) Activity recall in a visual cortical ensemble. *Nature Neuroscience* 15: 449–455. doi: [10.1038/nn.3036](https://doi.org/10.1038/nn.3036) PMID: [22267160](https://pubmed.ncbi.nlm.nih.gov/22267160/)
34. Davidson TJ, Kloosterman F, Wilson Ma (2009) Hippocampal Replay of Extended Experience. *Neuron* 63: 497–507. doi: [10.1016/j.neuron.2009.07.027](https://doi.org/10.1016/j.neuron.2009.07.027) PMID: [19709631](https://pubmed.ncbi.nlm.nih.gov/19709631/)
35. Ziegler L, Zenke F, Kastner DB, Gerstner W (2015) Synaptic Consolidation: From Synapses to Behavioral Modeling. *Journal of Neuroscience* 35: 1319–1334. doi: [10.1523/JNEUROSCI.3989-14.2015](https://doi.org/10.1523/JNEUROSCI.3989-14.2015) PMID: [25609644](https://pubmed.ncbi.nlm.nih.gov/25609644/)
36. Clopath C, Ziegler L, Vasilaki E, Büsing L, Gerstner W (2008) Tag-Trigger-Consolidation: A Model of Early and Late Long-Term-Potential and Depression. *PLoS Computational Biology* 4: e1000248. doi: [10.1371/journal.pcbi.1000248](https://doi.org/10.1371/journal.pcbi.1000248) PMID: [19112486](https://pubmed.ncbi.nlm.nih.gov/19112486/)
37. Pfister JP, Gerstner W (2006) Triplets of spikes in a model of spike timing-dependent plasticity. *The Journal of Neuroscience* 26: 9673–82. doi: [10.1523/JNEUROSCI.1425-06.2006](https://doi.org/10.1523/JNEUROSCI.1425-06.2006) PMID: [16988038](https://pubmed.ncbi.nlm.nih.gov/16988038/)
38. Zhang JC, Lau PM, Bi GQ (2009) Gain in sensitivity and loss in temporal contrast of STDP by dopaminergic modulation at hippocampal synapses. *Proceedings of the National Academy of Sciences of the United States of America* 106: 13028–13033. doi: [10.1073/pnas.0900546106](https://doi.org/10.1073/pnas.0900546106) PMID: [19620735](https://pubmed.ncbi.nlm.nih.gov/19620735/)
39. Kolodziejcki C, Porr B, Wörgötter F (2009) On the asymptotic equivalence between differential Hebbian and temporal difference learning. *Neural Computation* 21: 1173–1202. doi: [10.1162/neco.2008.04-08-750](https://doi.org/10.1162/neco.2008.04-08-750) PMID: [19018698](https://pubmed.ncbi.nlm.nih.gov/19018698/)
40. Potjans W, Morrison A, Diesmann M (2009) A Spiking Neural Network Model of an Actor-Critic Agent. *Neural Computation* 21: 301–339. doi: [10.1162/neco.2008.08-07-593](https://doi.org/10.1162/neco.2008.08-07-593) PMID: [19196231](https://pubmed.ncbi.nlm.nih.gov/19196231/)
41. Frémaux N, Sprekeler H, Gerstner W (2013) Reinforcement learning using a continuous time actor-critic framework with spiking neurons. *PLoS computational biology* 9: e1003024. doi: [10.1371/journal.pcbi.1003024](https://doi.org/10.1371/journal.pcbi.1003024) PMID: [23592970](https://pubmed.ncbi.nlm.nih.gov/23592970/)
42. Maass W, Natschläger T, Markram H (2002) Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Computation* 14: 2531–60. doi: [10.1162/089976602760407955](https://doi.org/10.1162/089976602760407955) PMID: [12433288](https://pubmed.ncbi.nlm.nih.gov/12433288/)
43. Jaeger H, Haas H (2004) Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* 304: 78–80. doi: [10.1126/science.1091277](https://doi.org/10.1126/science.1091277) PMID: [15064413](https://pubmed.ncbi.nlm.nih.gov/15064413/)
44. Drew PJ, Abbott LF (2006) Extending the effects of spike-timing-dependent plasticity to behavioral timescales. *Proceedings of the National Academy of Sciences of the United States of America* 103: 8876–81. doi: [10.1073/pnas.0600676103](https://doi.org/10.1073/pnas.0600676103) PMID: [16731625](https://pubmed.ncbi.nlm.nih.gov/16731625/)
45. Huertas MA, Shuler XMGH, Shouval HZ (2015) A Simple Network Architecture Accounts for Diverse Reward Time Responses in Primary Visual Cortex 35: 12659–12672.
46. Latimer KW, Yates JL, Meister MLR, Huk AC, Pillow JW (2015) Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science* 349: 184–187. doi: [10.1126/science.aaa4056](https://doi.org/10.1126/science.aaa4056) PMID: [26160947](https://pubmed.ncbi.nlm.nih.gov/26160947/)
47. Palmer SE, Marre O, Berry MJ, Bialek W (2015) Predictive information in a sensory population. *PNAS* 112:6908–13. doi: [10.1073/pnas.1506855112](https://doi.org/10.1073/pnas.1506855112) PMID: [26038544](https://pubmed.ncbi.nlm.nih.gov/26038544/)
48. Kushner HJ, Yin GG (2003) *Stochastic Approximation and Recursive Algorithms and Applications*, volume 35. Springer Science & Business Media.