LETTER Communicated by Laurence Abbott

Spike-Timing-Dependent Hebbian Plasticity as Temporal Difference Learning

Rajesh P. N. Rao
Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195-2350, U.S.A.

Terrence J. Sejnowski
Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, CA 92037, U.S.A., and Department of Biology, University of California at San Diego, La Jolla, CA 92037, U.S.A.

A spike-timing-dependent Hebbian mechanism governs the plasticity of recurrent excitatory synapses in the neocortex: synapses that are activated a few milliseconds before a postsynaptic spike are potentiated, while those that are activated a few milliseconds after are depressed. We show that such a mechanism can implement a form of temporal difference learning for prediction of input sequences. Using a biophysical model of a cortical neuron, we show that a temporal difference rule used in conjunction with dendritic backpropagating action potentials reproduces the temporally asymmetric window of Hebbian plasticity observed physiologically. Furthermore, the size and shape of the window vary with the distance of the synapse from the soma. Using a simple example, we show how a spike-timing-based temporal difference learning rule can allow a network of neocortical neurons to predict an input a few milliseconds before the input's expected arrival.

1 Introduction

Neocortical circuits are dominated by massive excitatory feedback: more than 80% of the synapses made by excitatory cortical neurons are onto other excitatory cortical neurons (Douglas, Koch, Mahowald, Martin, & Suarez, 1995). Why is there such massive recurrent excitation in the neocortex, and what is its role in cortical computation? Previous modeling studies have suggested a role for excitatory feedback in amplifying feedforward inputs (Douglas et al., 1995; Suarez, Koch, & Douglas, 1995; Ben-Yishai, Bar-Or, & Sompolinsky, 1995; Somers, Nelson, & Sur, 1995; Chance, Nelson, & Abbott, 1999). Recently, however, it has been shown that recurrent excitatory connections between cortical neurons are modified according to a spike-timing-dependent temporally asymmetric Hebbian learning rule: synapses that are activated slightly before the cell fires are strengthened whereas those that

are activated slightly after are weakened (Levy & Steward, 1983; Markram, Lubke, Frotscher, & Sakmann, 1997; Gerstner, Kempter, van Hemmen, & Wagner, 1996; Zhang, Tao, Holt, Harris, & Poo, 1998; Bi & Poo, 1998; Sejnowski, 1999). Information regarding the postsynaptic activity of the cell is conveyed back to the dendritic locations of synapses by backpropagating action potentials from the soma (Stuart & Sakmann, 1994).

In this article, we explore the hypothesis that recurrent excitation in neocortical circuits subserves the function of prediction and generation of temporal sequences (for related ideas, see Jordan, 1986; Elman, 1990; Minai & Levy, 1993; Montague & Sejonowski, 1994; Abbott & Blum, 1996; Rao & Ballard, 1997; Barlow, 1998; Westerman, Northmore, & Elias, 1999). In particular, we show that a temporal-difference-based learning rule for prediction (Sutton, 1988), when applied to backpropagating action potentials in dendrites, reproduces the temporally asymmetric window of Hebbian plasticity obtained in physiological experiments (see section 3). We examine the stability of the learning rule in section 4 and discuss possible biophysical mechanisms for implementing this rule in section 5. We also provide a simple example demonstrating how such a learning mechanism may allow cortical networks to learn to predict their inputs using recurrent excitation. The model predicts that cortical neurons may employ different temporal windows of plasticity at different dendritic locations to allow them to capture correlations between pre- and postsynaptic activity at different timescales (see section 6). A preliminary report of this work appeared as Rao and Sejnowski (2000).

2 Temporal Difference Learning

To predict input sequences accurately, the recurrent excitatory connections in a given network need to be adjusted such that the appropriate set of neurons is activated at each time step. This can be achieved by using a temporal difference (TD) learning rule (Sutton, 1988; Montague & Sejnowski, 1994). In this paradigm of synaptic plasticity, an activated synapse is strengthened or weakened based on whether the difference between two temporally separated predictions is positive or negative. This minimizes the errors in prediction by ensuring that the prediction generated by the neuron after synaptic modification is closer to the desired value than before.

The simplest example of a TD learning rule arises in the problem of predicting a scalar quantity z using a neuron with synaptic weights $w(1), \ldots, w(k)$ (represented as a vector $\mathbf{w}$). The neuron receives as presynaptic input the sequence of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_m$. The output of the neuron at time t is assumed to be given by $P_t = \sum_i w(i)x_t(i)$. The goal is to learn a set of synaptic weights such that the prediction P_t is as close as possible to the target z. According to the temporal difference (TD(0)) learning rule (Sutton, 1988), the weights are changed at time t by an amount given by:

$$\Delta \mathbf{w}_t = \alpha(P_{t+1} - P_t)\mathbf{x}_t, \tag{2.1}$$

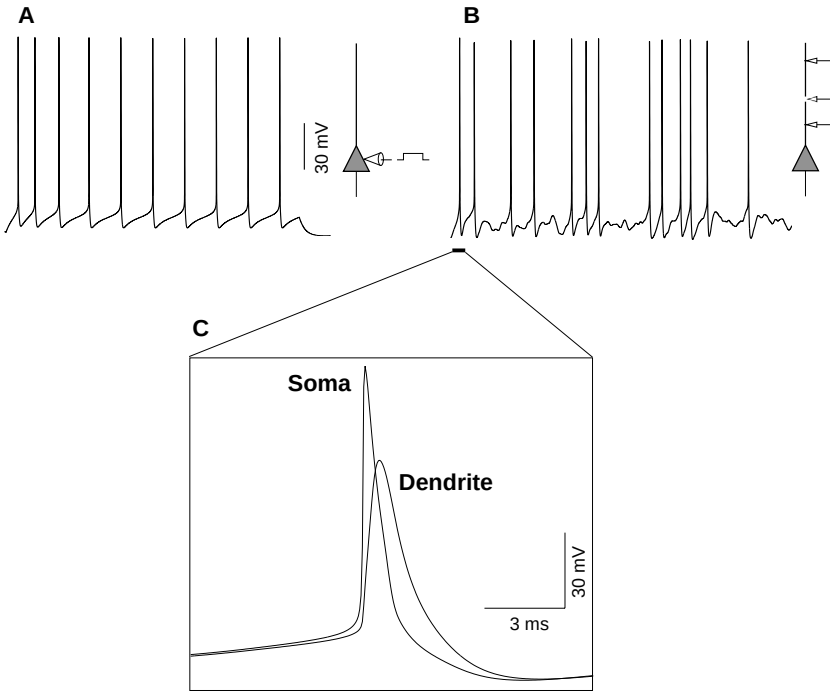

Figure 1: Model neuron response properties. (A) Response of a model neuron to a 70 pA current pulse injection into the soma for 900 milliseconds. (B) Response of the same model neuron to Poisson distributed excitatory and inhibitory synaptic inputs at random locations on the dendrite. (C) Example of a backpropagating action potential in the dendrite of the model neuron as compared to the corresponding action potential in the soma (enlarged from the initial portion of the trace in B).

where α is a learning rate or gain parameter and $P_{m+1} = z$. Note that in such a learning paradigm, synaptic plasticity is governed by the TD in postsynaptic activity at time instants $t + 1$ and t in conjunction with presynaptic activity $\mathbf{x}_t$ at time t. We use this paradigm of learning in the next section to model spike-timing-dependent plasticity.

3 Modeling Spike-Timing-Dependent Plasticity as TD Learning

In order to ascertain whether spike-timing-dependent temporally asymmetric plasticity in cortical neurons can be interpreted as a form of TD learning, we used a two-compartment model of a cortical neuron consisting of a dendrite and a soma-axon compartment (see Figure 1). The compartmental model was based on a previous study that demonstrated the

ability of such a model to reproduce a range of cortical response properties (Mainen & Sejnowski, 1996). Four voltage-dependent currents and one calcium-dependent current were simulated (as in Mainen & Sejnowski, 1996): fast Na^+, I_{Na}; fast K^+, I_{Kv}; slow noninactivating K^+, I_{Km}; high voltage-activated Ca^{2+}, I_{Ca}; and calcium-dependent K^+ current, I_{KCa}. The following active conductance densities were used in the soma-axon compartment (in $pS/\mu m^2$): $\overline{g}_{Na} = 40{,}000$ and $\overline{g}_{Kv} = 1400$. For the dendritic compartment, we used: $\overline{g}_{Na} = 20$, $\overline{g}_{Ca} = 0.2$, $\overline{g}_{Km} = 0.1$, and $\overline{g}_{KCa} = 3$, with leak conductance 33.3 $\mu S/cm^2$ and specific membrane resistance 30 $k\Omega$-cm^2. The presence of voltage-activated sodium channels in the dendrite allowed backpropagation of action potentials from the soma into the dendrite as shown in Figure 1C.

Conventional Hodgkin-Huxley-type kinetics were used for all currents (integration time step = 25 μs, temperature = 37°C). Ionic currents I were calculated using the ohmic equation,

$$I = \overline{g} A^x B (V - E), \tag{3.1}$$

where $\overline{g}$ is the maximal ionic conductance density, A and B are activation and inactivation variables, respectively (x denotes the order of kinetics—see Mainen & Sejonowski, 1996), and E is the reversal potential for the given ion species ($E_K = -90$ mV, $E_{Na} = 60$ mV, $E_{Ca} = 140$ mV, $E_{leak} = -70$ mV). For all compartments, the specific membrane capacitance was 0.75 $\mu F/cm^2$. Two key parameters governing the response properties of the model neuron are (Mainen & Sejnowski, 1996): the ratio of axo-somatic area to dendritic membrane area (ρ) and the coupling resistance between the two compartments (κ). For the simulations, we used the values $\rho = 150$ (with an area of 100 μm^2 for the soma-axon compartment) and a coupling resistance of $\kappa = 8\ M\Omega$. Poisson-distributed synaptic inputs to the dendrite (see Figure 1B) were simulated using alpha function–shaped (Koch, 1999) current pulse injections (time constant = 5 ms) at Poisson intervals with a mean presynaptic firing frequency of 3 Hz.

To study plasticity, excitatory postsynaptic potentials (EPSPs) were elicited at different time delays with respect to postsynaptic spiking by presynaptic activation of a single excitatory synapse located on the dendrite. Synaptic currents were calculated using a kinetic model of synaptic transmission with model parameters fitted to whole-cell recorded AMPA currents (see Destexhe, Mainen, & Sejnowski, 1998 for more details). Synaptic plasticity was simulated by incrementing or decrementing the value for maximal synaptic conductance by an amount proportional to the TD in the postsynaptic membrane potential at time instants $t + \Delta t$ and t for presynaptic activation at time t. The delay parameter Δt was set to 10 ms to yield results consistent with previous physiological experiments (Markram et al., 1997; Bi & Poo, 1998). Presynaptic input to the model neuron was paired with postsynaptic spiking by injecting a depolarizing current pulse (10 ms, 200 pA) into

the soma. Changes in synaptic efficacy were monitored by applying a test stimulus before and after pairing and recording the EPSP evoked by the test stimulus.

Figure 2 shows the results of pairings in which the postsynaptic spike was triggered 5 ms after and 5 ms before the onset of the EPSP, respectively. While the peak EPSP amplitude was increased 58.5% in the former case, it was decreased 49.4% in the latter case, qualitatively similar to experimental observations (Markram et al., 1997; Bi & Poo, 1998). The critical window for synaptic modifications in the model depends on the parameter Δt as well as the shape of the backpropagating action potential. This window of plasticity was examined by varying the time interval between presynaptic stimulation and postsynaptic spiking (with $\Delta t = 10$ ms). As shown in Figure 2C, changes in synaptic efficacy exhibited a highly asymmetric dependence on spike timing similar to physiological data (Markram et al., 1997). Potentiation was observed for EPSPs that occurred between 1 and 12 ms before the postsynaptic spike, with maximal potentiation at 6 ms. Maximal depression was observed for EPSPs occurring 6 ms after the peak of the postsynaptic spike, and this depression gradually decreased, approaching zero for delays greater than 10 ms. As in rat neocortical neurons (Markram et al., 1997), *Xenopus* tectal neurons (Zhang et al, 1998), and cultured hippocampal neurons (Bi & Poo, 1998), a narrow transition zone (roughly 3 ms in the model) separated the potentiation and depression windows.

4 Stability of the Learning Rule

A crucial question regarding the spike-based Hebbian learning rule described above is whether it produces a stable set of weights for a given training set of inputs. In the case of the conventional Hebbian learning rule, which only prescribes increases in synaptic weights based on pre- and postsynaptic correlations, numerous methods have been suggested to ensure stability, such as weight normalization and weight decay (see Sejnowski, 1977, and Montague & Sejnowski, 1994, for reviews). The classical TD learning rule is self-stabilizing because weight modifications are dependent on the error in prediction ($P_{t+1} - P_t$), which can be positive or negative. Thus, when the weights $\mathbf{w}$ become large, the predictions P_t tend to become large compared to P_{t+1} (e.g., at the end of the training sequence). This results in a decrease in synaptic strength, as needed to ensure stability.

In the case of the spike-based learning rule, convergence to a stable set of weights depends on both the biophysical properties of the neuron and the TD parameter Δt. To see this, consider the example in Figure 3, where a presynaptic spike occurs at time t_{pre} and causes a backpropagating action potential (AP) at time t_{BPAP} in a postsynaptic model neuron (model neuron parameters were the same as those given in section 3). Using a fixed $\Delta t = 10$ milliseconds, the synaptic conductance initially increases because the TD error ($P_{t_{pre}+\Delta t} - P_{t_{pre}}$) is positive. The increase in synaptic conduc-

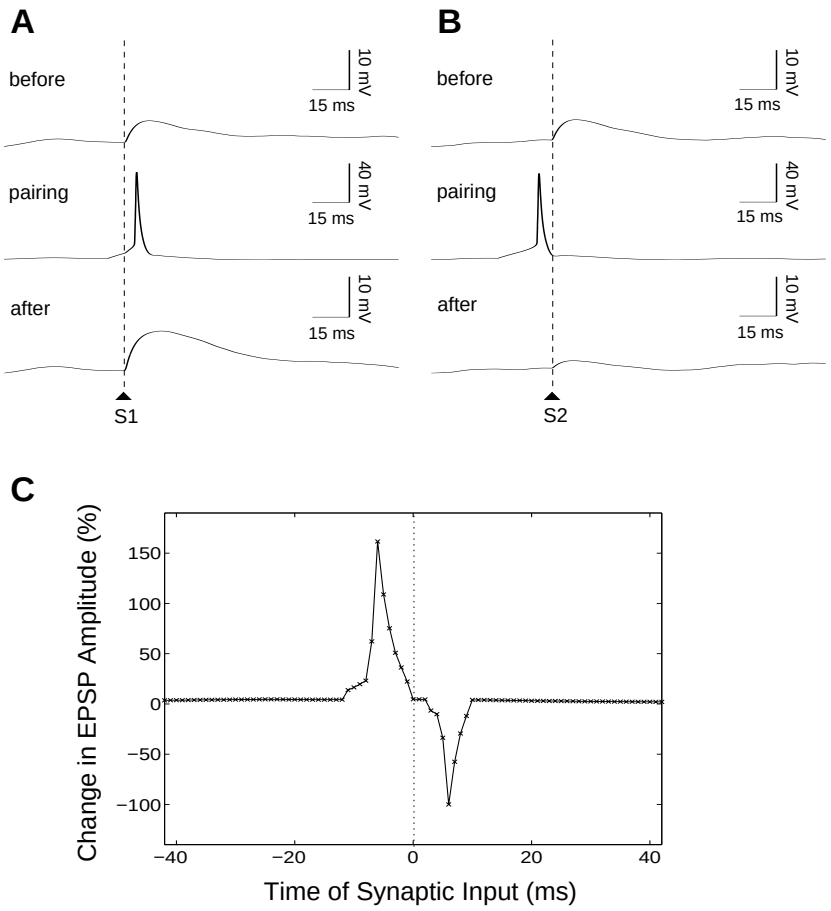

Figure 2: Synaptic plasticity in a model neocortical neuron. (A) EPSP in the model neuron evoked by a presynaptic spike (S1) at an excitatory synapse ("before"). Pairing this presynaptic spike with postsynaptic spiking after a 5 ms delay ("pairing") induces long-term potentiation ("after"). (B) If presynaptic stimulation (S2) occurs 5 ms after postsynaptic firing, the synapse is weakened, resulting in a corresponding decrease in peak EPSP amplitude. (C) Temporally asymmetric window of synaptic plasticity obtained by varying the delay between pre- and postsynaptic spiking (negative delays refer to presynaptic before postsynaptic spiking).

tance causes the postsynaptic neuron to fire earlier, which in turn causes an increase in the TD error, until $P_{t_{pre}+\Delta t}$ reaches a value close to $P_{t_{pre}}$. This happens when the time of the somatic AP is almost equal to t_{pre} and the

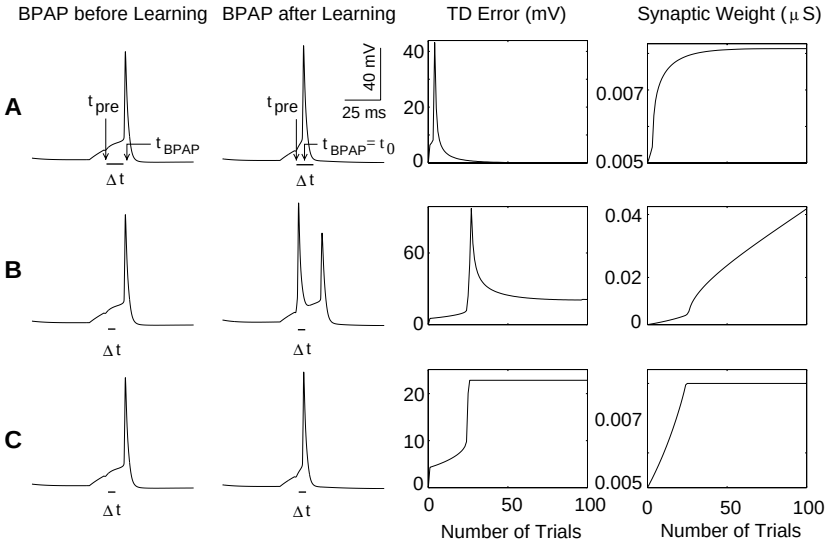

Figure 3: Stability of spike-timing based temporal difference learning. (A) A presynaptic spike at time t_{pre} results in a backpropagating action potential (BPAP) a few milliseconds later at time t_{BPAP}, causing an increase in synaptic strength due to spike-timing-dependent plasticity. The TD parameter Δt was set to 10 milliseconds. After 100 trials of learning (pre- and postsynaptic pairing), the TD error is reduced to zero, and the maximal synaptic conductance converges to a stable value of 8.1 nS. The postsynaptic BPAP now occurs at an earlier time t_0 compared to its time of occurrence before learning. (B) When a shorter Δt of 5 ms is used (with the same initial conditions as in A), the synaptic strength continues to grow without bound after 100 trials because the TD error remains positive. A burst of two spikes is elicited due to an unrealistically large synaptic conductance. (C) An upper bound of 8 nS on the maximal synaptic conductance prevents unbounded synaptic growth and ensures that a single spike is elicited.

backpropagating AP occurs slightly later at time $t_{BPAP} = t_0$ (Figure 3A, second graph). The synaptic conductance then remains stable for this pair of pre- and postsynaptic spikes (Figure 3A, right-most graph).

On the other hand, if Δt is much smaller than the width of the backpropagating AP, the maximal synaptic conductance may grow without bound, as shown in Figure 3B (right-most graph). In this case ($\Delta t = 5$ ms), the backpropagating AP does not peak soon enough after the presynaptic spike, causing the TD error to remain positive. As a result, the synaptic conductance does not converge to a stable value, and multiple spikes, in the form of a burst, are elicited (see Figure 3B, second graph). One solution to the problem of unbounded growth in synaptic conductance is to place an up-

per bound on the synaptic strength, as suggested by Abbott and Song (1999) and Gerstner et al. (1996). In this case, although the TD error remains positive, the synaptic conductance remains clamped at the upper bound value, and only a single spike is elicited (see Figure 3C). The use of an upper bound is partially supported by physiological experiments showing a dependence of spike-timing-based synaptic plasticity on initial synaptic strength: long-term potentiation (LTP) was found to occur mostly in weak synapses, while connections that already had large values for their synaptic strength consistently failed to show LTP (Bi & Poo, 1998).

In summary, the stability of the TD learning rule for spike-timing-dependent synaptic plasticity depends crucially on whether the temporal window parameter Δt is comparable in magnitude to the width of the backpropagating AP at the location of the synapse. An upper bound on the maximal synaptic conductance may be required to ensure stability in general (Gerstner et al., 1996; Abbott & Song, 1999; Song, Miller, & Abbott, 2000). Such a saturation constraint is partially supported by experimental data (Bi & Poo, 1998). An alternative approach that might be worth considering is to explore a learning rule that uses a continuous form of the TD error, where, for example, an average of postsynaptic activity is subtracted from the current activity (Montague & Sejnowski, 1994; Doya, 2000). Such a rule may offer better stability properties than the discrete TD rule that we have used, although other parameters, such as the window over which average activity is computed, may still need to be chosen carefully.

5 Biophysical Mechanisms

An interesting question is whether a biophysical basis can be found for the TD learning model described above. Neurophysiological and imaging studies suggest a role for dendritic Ca^{2+} signals in the induction of spike-timing-dependent LTP and LTD (long-term depression) in hippocampal and cortical neurons (Magee & Johnston, 1997; Koester & Sakmann, 1998; Paulsen & Sejnowski, 2000). In particular, when an EPSP preceded a postsynaptic action potential, the Ca^{2+} transient in dendritic spines, where most excitatory synaptic connections occur, was observed to be larger than the sum of the Ca^{2+} signals generated by the EPSP or AP alone, causing LTP; on the other hand, when the EPSP occurred after the AP, the Ca^{2+} transient was found to be a sublinear sum of the signals generated by the EPSP or AP alone, resulting in LTD (Koester & Sakmann, 1998; Linden, 1999; Paulsen & Sejnowski, 2000). Possible sources contributing to the spinous Ca^{2+} transients include Ca^{2+} ions entering through NMDA receptors (Bliss & Collingridge, 1993; Koester & Sakmann, 1998), voltage-gated Ca^{2+} channels in the dendrites (Schiller, Schiller, & Clapham, 1998), and calcium-induced calcium release from intracellular stores (Emptage, 1999).

How do the above experimental observations support the TD model? In a recent study (Franks, Bartol, Egelman, Poo, & Sejnowski, 1999), a Monte

Carlo simulation program MCell (Stiles, Bartol, Salpeter, Salpeter, & Sejnowski, 2000) was used to model the Ca^{2+} dynamics in dendritic spines following pre- and postsynaptic activity and to track the binding of Ca^{2+} to endogenous proteins. The influx of Ca^{2+} into a spine is governed by the rapid depolarization pulse caused by the backpropagating AP. The width of the backpropagating AP is much smaller than the time course of glutamate binding to the NMDA receptor. As a result, the dynamics of Ca^{2+} influx and binding to calcium-binding proteins such as calmodulin depends highly nonlinearly on the relative timing of presynaptic activation (with release of glutamate) and postsynaptic depolarization (due to the backpropagating AP). In particular, due to its kinetics, the binding protein calmodulin could serve as a differentiator of intracellular calcium concentration, causing synapses either to potentiate or depress depending on the spatiotemporal profile of the dendritic Ca^{2+} signal (Franks et al., 1999). As a consequence of these biophysical mechanisms, the change in synaptic strength depends, to a first approximation, on the time derivative of the postsynaptic activity, as postulated by the TD model.

6 Model Predictions

The shape and size of backpropagating APs at different locations in a cortical dendrite depend on the distance of the dendritic location from the soma. For example, as backpropagating APs progress from the soma to the distal parts of a dendrite, they tend to become broader as compared to more proximal parts of the dendrite. This is shown in Figure 4 for a reconstructed layer 5 neocortical neuron (Douglas, Martin, & Whitteridge, 1991) from cat visual cortex with ionic channels based on those used for this neuron in the study by Mainen and Sejnowski (1996).

Since synaptic plasticity in our model depends on the TD in postsynaptic activity, the model predicts that synapses situated at different locations on a dendrite should exhibit different temporally asymmetric windows of plasticity. This is illustrated in Figure 4 for the reconstructed model neuron. Learning windows were calculated by applying the TD operator to the backpropagating APs with $\Delta t = 2$ ms. The model predicts that the window of plasticity for distal synapses should be broader than for proximal synapses. Having broader windows would allow distal synapses to encode longer timescale correlations between pre- and postsynaptic activity. Proximal synapses would encode correlations at shorter timescales due to sharper learning windows. Thus, by distributing its synapses throughout its dendritic tree, a cortical neuron could in principle capture a wide range of temporal correlations between its inputs and its output. This in turn would allow a network of cortical neurons to predict sequences accurately and reduce possible ambiguities such as aliasing between learned sequences by tracking the sequence at multiple timescales (Rao & Ballard, 1997; Rao, 1999).

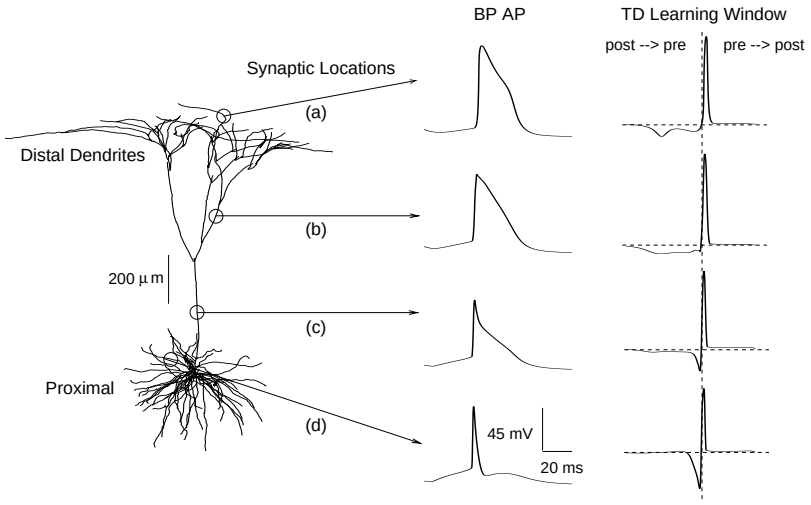

Figure 4: Dependence of learning window on synaptic location. The traces in the middle column show size and shape of a backpropagating AP at different dendritic locations in a compartmental model of a reconstructed layer 5 neocortical neuron. The corresponding TD learning window for putative synapses at these dendritic locations is shown on the right. Note the gradual broadening of the learning window in time as one progresses from proximal to distal synapses.

The TD model also predicts asymmetries in size and shape between the LTP and LTD windows. For example, in Figure 4, the LTD window for the two apical synapses (labeled a and b) is much broader and shallower than the corresponding LTP window. Such an asymmetry between LTP and LTD has recently been reported for synapses in rat primary somatosensory cortex (S1) (Feldman, 2000). In particular, the range of time delays between pre- and postsynaptic spiking that induces LTD was found to be much longer than the range of delays that induces LTP, generating learning windows similar to the top two windows in Figure 4. One computational consequence of such a learning window is that synapses that elicit subthreshold EPSPs in a manner uncorrelated with postsynaptic spiking will become depressed over time. In rat primary somatosensory cortex, plucking one whisker but sparing its neighbor causes neuronal responses to the deprived whisker in layer II/III to become rapidly depressed. The asymmetry in the LTP/LTD learning windows provides an explanation for this phenomenon: spontaneously spiking inputs from plucked whiskers are uncorrelated with postsynaptic spiking, and therefore, synapses receiving these inputs will become depressed (Feldman, 2000). Such a mechanism may contribute to experience-dependent depression of responses and related changes in the receptive field properties of neurons in other cortical areas as well.

7 Discussion

Our results suggest that spike-timing-dependent plasticity in neocortical synapses can be interpreted as a form of TD learning for prediction. To see how a network of model neurons can learn to predict sequences using such a learning mechanism, consider the simple case of two excitatory neurons N1 and N2 connected to each other, receiving inputs from two separate input neurons I1 and I2 (see Figure 5A). Model neuron parameters were the same as those used in section 3. Suppose input neuron I1 fires before input neuron I2, causing neuron N1 to fire (see Figure 5B). The spike from N1 results in a subthreshold EPSP in N2 due to the synapse S2. If input arrives from I2 between 1 and 12 ms after this EPSP and if the temporal summation of these two EPSPs causes N2 to fire, synapse S2 will be strengthened. The synapse S1, on the other hand, will be weakened because the EPSP due to N2 arrives a few milliseconds after N1 has fired.

After several exposures to the I1–I2 training sequence, when I1 causes neuron N1 to fire, N1 in turn causes N2 to fire several milliseconds *before* input I2 occurs due to the potentiation of the recurrent synapse S2 in previous trials (see Figure 5C). Input neuron I2 can thus be inhibited by the predictive feedback from N2 just before the occurrence of imminent input activity (marked by the asterisk in Figure 5C). This inhibition prevents input I2 from further exciting N2, thereby implementing a negative feedback–based predictive coding circuit (Rao & Ballard, 1999). Similarly, a positive feedback loop between neurons N1 and N2 is avoided because the synapse S1 was weakened in previous trials (see the arrows in Figures 5B and 5C, top row). Figure 6A depicts the process of potentiation and depression of the two synapses as a function of the number of exposures to the I1–I2 input sequence. The decrease in latency of the predictive spike elicited in N2 with respect to the timing of input I2 is shown in Figure 6B. Notice that before learning, the spike occurs 3.2 ms after the occurrence of the input, whereas after learning, it occurs 7.7 ms before the input. Although the postsynaptic spike continues to occur shortly after the activation of synapse S2, this synapse is prevented from assuming larger values due to a saturation constraint of 0.03 μS on the maximal synaptic conductance (see section 4 for a discussion of this constraint). In related work (Rao & Sejnowkski, 2000), we have shown how such a learning mechanism can explain the development of direction selectivity in recurrent cortical networks, yielding receptive field properties similar to those observed in awake monkey V1.

The precise biophysical mechanisms underlying spike-timing-dependent temporal difference learning remain unclear. However, as discussed in section 5, calcium fluctuations in dendritic spines are known to be strongly dependent on the timing between pre- and postsynaptic spikes. Such calcium transients may cause, via calcium-mediated signaling cascades, asymmetric synaptic modifications that are dependent, to a first approximation,

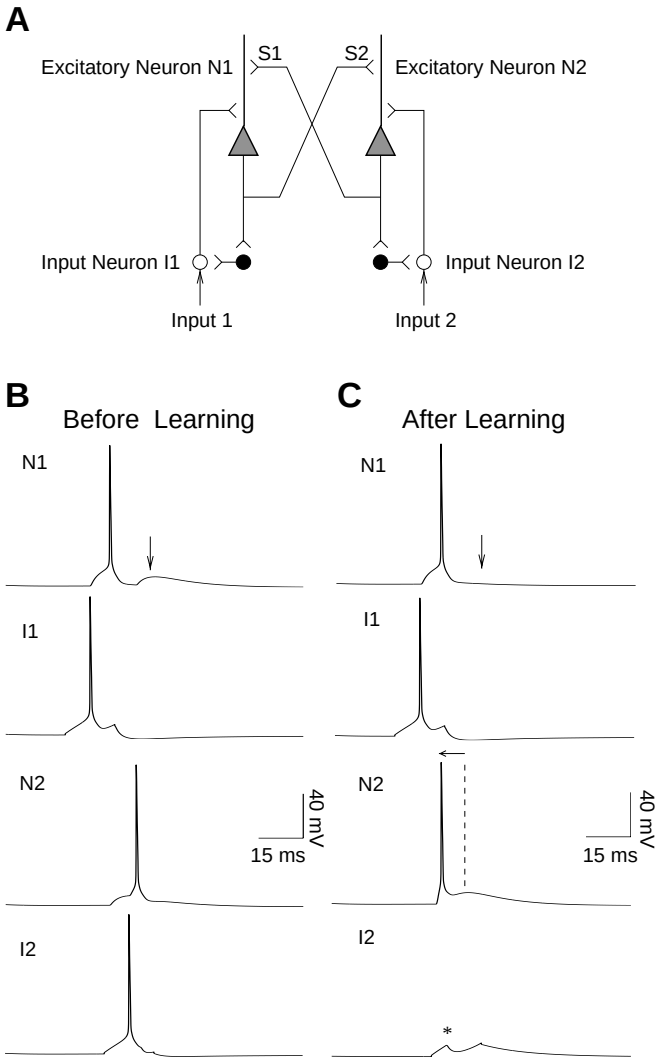

Figure 5: Learning to predict using spike-timing-dependent Hebbian learning. (A) Network of two model neurons N1 and N2 recurrently connected via excitatory synapses S1 and S2, with input neurons I1 and I2. N1 and N2 inhibit the input neurons via inhibitory interneurons (dark circles). (B) Network activity elicited by the sequence I1 followed by I2. (C) Network activity for the same sequence after 40 trials of learning. Due to strengthening of recurrent synapse S2, recurrent excitation from N1 now causes N2 to fire several milliseconds before the expected arrival of input I2 (dashed line), allowing it to inhibit I2 (asterisk). Synapse S1 has been weakened, preventing reexcitation of N1 (downward arrows show a decrease in EPSP).

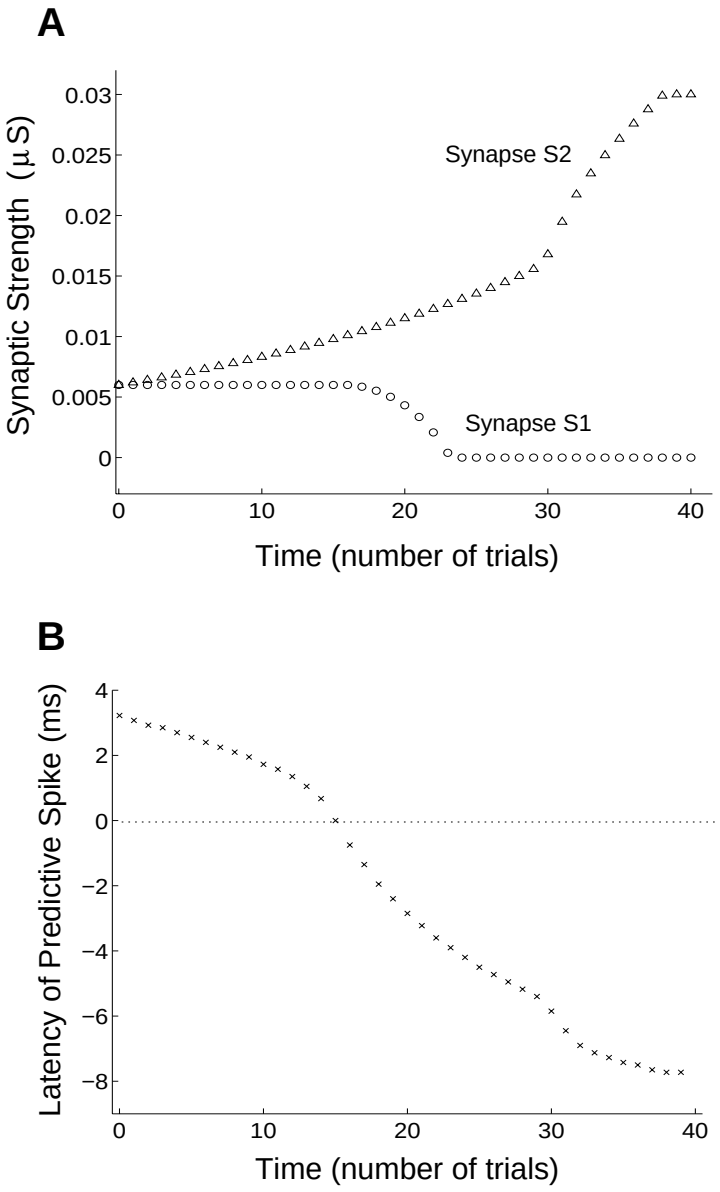

Figure 6: Synaptic strength and latency reduction due to learning. (A) Potentiation and depression of synapses S1 and S2, respectively, during the course of learning. Synaptic strength was defined as maximal synaptic conductance in the kinetic model of synaptic transmission (Destexhe et al., 1998). (B) Latency of predictive spike in N2 during the course of learning measured with respect to the time of input spike in I2 (dotted line).

on the temporal derivative of postsynaptic activity. An interesting topic worthy of further investigation is therefore the development of more realistic implementations of TD learning based on, for instance, the temporal derivative of postsynaptic calcium activity rather than the TD in postsynaptic membrane potential as modeled here.

An alternate approach to analyzing spike-timing-dependent learning rules is to decompose an observed asymmetric learning window into a TD component plus noise and to analyze the noise component. However, the advantage of our approach is that one can predict the shape of plasticity windows at different dendritic locations based on an estimate of postsynaptic activity, as described in section 6. Conversely, given a particular learning window, one can use the model to explain the temporal asymmetry of the window as a function of the neuron's postsynaptic activity profile.

Temporally asymmetric learning has previously been suggested as a possible mechanism for sequence learning in the hippocampus (Minai & Levy, 1993; Abbott & Blum, 1996) and as an explanation for the asymmetric expansion of hippocampal place fields during route learning (Mehta, Barnes, & McNaughton, 1997). Some of these models used relatively long temporal windows of synaptic plasticity, on the order of several hundreds of milliseconds (Abbott & Blum, 1996), while others used temporal windows in the submillisecond range for coincidence detection (Gerstner et al., 1996). Sequence learning in our model is based on a window of plasticity that spans approximately ± 20 milliseconds, which is roughly consistent with recent physiological observations (Markram et al., 1997) (see also Abbott & Song, 1999; Roberts, 1999; Westerman et al., 1999; Mehta & Wilson, 2000; Song et al., 2000).

Several theories of prediction and sequence learning in the hippocampus and neocortex have been proposed based on statistical and information theoretic ideas (Minai & Levy, 1993; Montague & Sejnowski, 1994; Abbott & Blum, 1996; Dayan & Hinton, 1996; Daugman & Downing, 1995; Barlow, 1998; Rao & Ballard, 1999). Our biophysical simulations suggest a possible implementation of such models in cortical circuitry. Given the universality of the problem of encoding and generating temporal sequences in both sensory and motor domains, the hypothesis of TD-based sequence learning in recurrent neocortical circuits may help provide a unifying principle for studying cortical structure and function.

Acknowledgments

This research was supported by the Sloan Center for Theoretical Neurobiology at the Salk Institute and the Howard Hughes Medical Institute. We thank Peter Dayan and an anonymous referee for their constructive comments and suggestions.

References

Abbott, L. F., & Blum, K. I. (1996). Functional significance of long-term potentiation for sequence learning and prediction. *Cereb. Cortex, 6*, 406–416.

Abbott, L. F., & Song, S. (1999). Temporally asymmetric Hebbian learning, spike timing and neural response variability. In M. S. Kearns, S. Solla, & D. Cohn (Eds.), *Advances in neural information processing systems, 11* (pp. 69–75). Cambridge, MA: MIT Press.

Barlow, H. (1998). Cerebral predictions. *Perception, 27*, 885–888.

Ben-Yishai, R., Bar-Or, R. L., & Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U.S.A., 92*, 3844–3848.

Bi, G. Q., & Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci., 18*, 10464–10472.

Bliss, T. V., & Collingridge, G. L. (1993). A synaptic model of memory: Long-term potentiation in the hippocampus. *Nature, 361*, 31–39.

Chance, F. S., Nelson, S. B., & Abbott, L. F. (1999). Complex cells as cortically amplified simple cells. *Nature Neuroscience, 2*, 277–282.

Daugman, J. G., & Downing, C. J. (1995). Demodulation, predictive coding, and spatial vision. *J. Opt. Soc. Am. A, 12*, 641–660.

Dayan, P., & Hinton, G. (1996). Varieties of Helmholtz machine. *Neural Networks, 9*(8), 1385–1403.

Destexhe, A., Mainen, Z., & Sejnowski, T. (1998). Kinetic models of synaptic transmission. In C. Koch & I. Segev (Eds.), *Methods in neuronal modeling*. Cambridge, MA: MIT Press.

Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A., & Suarez, H. H. (1995). Recurrent excitation in neocortical circuits. *Science, 269*, 981–985.

Douglas, R. J., Martin, K. A. C., & Whitteridge, D. (1991). An intracellular analysis of the visual responses of neurons in cat visual cortex. *J. Physiol., 440*, 659–696.

Doya, K. (2000). Reinforcement learning in continuous time and space. *Neural Computation, 12*(1), 219–245.

Elman, J. (1990). Finding structure in time. *Cognitive Science, 14*, 179–211.

Emptage, N. J. (1999). Calcium on the up: Supralinear calcium signaling in central neurons. *Neuron, 24*, 495–497.

Feldman, D. (2000). Timing-based LTP and LTD at vertical inputs to layer II/III pyramidal cells in rat barrel cortex. *Neuron, 27*, 45–56.

Franks, K. M., Bartol, T. M., Egelman, D. M., Poo, M. M., & Sejnowski, T. J. (1999). Simulated dendritic influx of calcium ions through voltage- and ligand-gated channels using MCell. *Soc. Neurosci. Abstr., 25*, 1989.

Gerstner, W., Kempter, R., van Hemmen, J. L., & Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. *Nature, 383*, 76–81.

Jordan, M. (1986). Attractor dynamics and parallelism in a connectionist sequential machine. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society* (pp. 531–546). Hillsdale, NJ: Erlbaum.

Koch, C. (1999). *Biophysics of computation: Information processing in single neurons*. New York: Oxford University Press.

Koester, H. J., & Sakmann, B. (1998). Calcium dynamics in single spines during coincident pre- and postsynaptic activity depend on relative timing of back-propagating action potentials and subthreshold excitatory postsynaptic potentials. *Proc. Natl. Acad. Sci. U.S.A., 95*(16), 9596–9601.

Levy, W. B., & Steward, O. (1983). Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience, 8*, 791–797.

Linden, D. J. (1999). The return of the spike: Postsynaptic action potentials and the induction of LTP and LTD. *Neuron, 22*(4), 661–666.

Magee, J. C., & Johnston, D. (1997). A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. *Science, 275*, 209–213.

Mainen, Z., & Sejnowski, T. (1996). Influence of dendritic structure on firing pattern in model neocortical neurons. *Nature, 382*, 363–366.

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

Mehta, M. R., Barnes, C. A., & McNaughton, B. L. (1997). Experience-dependent, asymmetric expansion of hippocampal place fields. *Proc. Natl. Acad. Sci. U.S.A., 94*, 8918–8921.

Mehta, M. R., & Wilson, M. (2000). From hippocampus to V1: Effect of LTP on spatiotemporal dynamics of receptive fields. *Neurocomputing, 32*, 905–911.

Minai, A. A., & Levy, W. B. (1993). Sequence learning in a single trial. In *Proceedings of the 1993 INNS World Congress on Neural Networks II* (pp. 505–508). Hillsdale, NJ: Erlbaum.

Montague, P. R., & Sejnowski, T. J. (1994). The predictive brain: Temporal coincidence and temporal order in synaptic learning mechanisms. *Learning and Memory, 1*, 1–33.

Paulsen, O., & Sejnowski, T. J. (2000). Natural patterns of activity and long-term synaptic plasticity. *Curr. Opin. Neurobiol., 10*(2), 172–179.

Rao, R. P. N. (1999). An optimal estimation approach to visual perception and learning. *Vision Research, 39*(11), 1963–1989.

Rao, R. P. N., & Ballard, D. H. (1997). Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Computation, 9*(4), 721–763.

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

Rao, R. P. N., & Sejnowski, T. J. (2000). Predictive sequence learning in recurrent neocortical circuits. In S. A. Solla, T. K. Leen, & K.-R. Muller (Eds.), *Advances in neural information processing systems, 12* (pp. 164–170). Cambridge, MA: MIT Press.

Roberts, P. D. (1999). Computational consequences of temporally asymmetric learning rules: I. Differential Hebbian learning. *J. Computational Neurosci., 7*, 235–246.

Schiller, J., Schiller, Y., & Clapham, D. E. (1998). NMDA receptors amplify calcium influx into dendritic spines during associative pre- and postsynaptic activation. *Nature Neurosci., 1*(2), 114–118.

Sejnowski, T. J. (1977). Storing covariance with nonlinearly interacting neurons. *Journal of Mathematical Biology, 4*, 203–211.

Sejnowski, T. J. (1999). The book of Hebb. *Neuron, 24*(4), 773–776.

Somers, D. C., Nelson, S. B., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci., 15*, 5448-5465.

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

Stiles, J. S., Bartol, T. M., Salpeter, M. M., Salpeter, E. E., & Sejnowski, T. J. (2000). Synaptic variability: New insights from reconstructions and Monte Carlo simulations with MCell. In M. Cowan & K. Davies (Eds.), *Synapses*. Baltimore: Johns Hopkins University Press.

Stuart, G. J., & Sakmann, B. (1994). Active propagation of somatic action potentials into neocortical pyramidal cell dendrites. *Nature, 367*, 69–72.

Suarez, H., Koch, C., & Douglas, R. (1995). Modeling direction selectivity of simple cells in striate visual cortex with the framework of the canonical microcircuit. *J. Neurosci., 15*(10), 6700–6719.

Sutton, R. S. (1988). Learning to predict by the method of temporal differences. *Machine Learning, 3*(1), 9–44.

Westerman, W. C., Northmore, D. P. M., & Elias, J. G. (1999). Antidromic spikes drive Hebbian learning in an artificial dendritic tree. *Analog Integrated Circuits and Signal Processing, 18*, 141–152.

Zhang, L. I., Tao, H. W., Holt, C. E., Harris, W. A., & Poo, M. M. (1998). A critical window for cooperation and competition among developing retinotectal synapses. *Nature, 395*, 37–44.

Received February 24, 2000; accepted December 27, 2000.