??, ??, 1–13 (??)

Computational Consequences of Temporally Asymmetric Learning Rules: I. Differential Hebbian Learning

PATRICK D. ROBERTS

proberts@reed.edu

Neurological Sciences Institute, OHSU, 1120 N.W. 20th Avenue, Portland, OR 97209

Received ??; Revised ??

Editors: ??

Abstract. Temporally asymetric learning rules governing plastic changes in synaptic efficacy have recently been identified in physiological studies. In these rules, the exact timing of pre- and postsynaptic spikes is critical to the induced change of synaptic efficacy. The temporal learning rules treated in this paper are approximately antisymmetric; the synaptic efficacy is enhanced if the postsynaptic spike follows the presynaptic spike by a few milliseconds, but the efficacy is depressed if the postsynaptic spike precedes the presynaptic spike. The learning dynamics of this rule are studied using a stochastic model neuron receiving a set of serially delayed inputs. The average change of synaptic efficacy due to the temporally antisymmetric learning rule is shown to yield differential Hebbian learning. These results are demonstrated with both mathematical analyses and computer simulations, and connections with theories of classical conditioning are discussed.

Keywords: synaptic plasticity, learning rule, classical conditioning, instability

1. Introduction

One of the great mysteries facing neuroscience today is how the nervous system learns causal relationships to form expectations of future events. A likely physiological basis of temporal associations would involve changes in the synaptic efficacy of connections among relevant neurons. Synaptic change has been suggested as the mechanism underlying the storage of memories in the nervous system [Hebb, 1949], and the recall of relevant temporal associations leads to biologically advantageous behavior.

Consistent correlation between two events may imply a causal connection, but more importantly, the correlation acts as a predictor of the later event based on the occurrence of the first. A neuronal representation of such a correlation would be when a neuron that is responding to a later event "learns" to respond to the earlier event following repeated pairing of the sequence. Eventually, there will be a neuronal response to the first event even in the absence of the second event. Continued repetition of the first event alone should reduce the neurons response by way of forgetting the learned temporal association. Thus, the predictive relations between events separated in time are maintained as long as they are reinforced.

A simple model of linking events in time by whole organisms is found in classical conditioning studies [Pavlov, 1927, Spence, 1956]. The response of an animal to a conditioned stimulus, after pairing with a later unconditioned stimulus, shows the ability to link the first stimulus to an anticipated occurrence of the second stimulus. Evidence of synaptic change in the amygdala has recently been observed during fear conditioning [McKernan and Shinnick-Gallagher, 1997, Rogan et al., 1997]. However, our understanding of the neurophysiological mechanisms underlying classical conditioning is as yet incomplete [Mauk, 1997].

Careful studies of associations between conditioned stimulus and reward have revealed a consistent pattern of conditioned behavior observed in the whole organism [Klopf, 1988, Sutton and Barto, 1981] that may be interpreted in terms of synaptic change. A learning rule based on the association of the *rate-of-change* of the pre- and postsynaptic activity, instead of simply levels of pre- and postsynaptic activity, yields results that closely match the learning curves observed behaviorally. This adjustment of synaptic efficacy by an amount proportional to a time derivative is termed *differential* Hebbian learning [Klopf, 1986, Kosko, 1986].

Some models have implemented differential Hebbian learning using comparisons between a time average of the neuronal activity and values [Montague et al., 1996]. Another class of models, called time-difference models, rely on a discrete version of the differential Hebbian algorithm by comparing neuronal activity between adjacent time steps [Sutton and Barto, 1981]. These models show promise in explaining many of the experimental classical conditioning paradigms and require synaptic mechanisms that would be physiologically similar to differential Hebbian models.

Models based on local rules of synaptic change are theoretically advantageous because they depend *only upon the neuronal activity at the time of association* so that all the information needed for synaptic change is available. Experimental investigations have confirmed that local changes in synaptic efficacy can be dependent on the precise temporal relations of pre- and postsynaptic spikes [Bell et al., 1997b, Levy and Steward, 1983, Markram et al., 1997b, Zhang et al., 1998]. These timing relations yield a local *temporal learning rule*.

In these types of learning rules, the presence and direction of synaptic change depends on the precise timing of pre- and postsynaptic spikes during the association period. The asymmetry is thought to result from the dynamics of long-term potentiation (or depression) mediated by glutamate receptors of the NMDA-type [Debanne et al., 1994, Gustafsson et al., 1987, Levy and Steward, 1983]. Because these receptor require the binding of glutamate in conjunction with postsynaptic depolarization for the influx of Ca^{2+} ions, the associated changes in synaptic efficacy occur if the postsynaptic depolarization *follows* the beginning of the epsp. Such asymetric learning rules has been previously investigated theoretically in the context of the hippocampal place cells [Abbott and Blum, 1996] and for temporal pattern recognition [Gerstner et al., 1993]. Measurements of the precise timing relationships have been recently demonstrated physiologically in the mammalian neocortex [Markram et al., 1997b], in a cerebellum-like structure of electric fish [Bell et al., 1997b], in the developing *Xenopus* optic tectum [Zhang et al., 1998], and in cultured hippocampal neurons [Bi and ming Poo, 1998].

The present study focuses on synapses between pyramidal cells of the rat neocortex that are enhanced or depressed depending on the timing of pairing pre- and postsynaptic spikes [Markram et al. 1997b] [Bi and ming Poo, 1998]. Evidence of postsynaptic spike propagation into the dendrites [Stuart and Sakmann, 1994] is thought to explain how postsynaptic activity is communicated to the site of the synapse. If the postsynaptic spike is 10 msec before the arrival of the presynaptic spike, then there is a depression of the synaptic efficacy. If the postsynaptic spike follows the arrival of the presynaptic spike by 10 msec, then there is an enhancement of synaptic efficacy (Fig. 3 in Markram et al. 1997b). No long term change in synaptic efficacy is observed at large time differences or in controls where there is no postsynaptic activity. A more complete characterization of an approximately antisymmetric learning rule has been made in the optic tectum [Zhang et al., 1998] and hippocampus [Bi and ming Poo, 1998]. In the following, we show that this learning rule,

that is predominantly antisymmetric in time, implies the presence of a differential Hebbian learning rule at synapses between pyramidal cells. In addition, due to the associative depression, this learning rule results in active forgetting of previously learned associations, an important aspect that was missing in previous theoretical studies [Abbott and Blum, 1996, Gerstner et al., l993].

Another type of asymmetric temporal learning rule that leads to an entirely different dynamics has been found in a cerebellum-like structure (the electrosensory lateral line lobe) in mormyrid electric fish. Purkinje-like cells in this structure exhibit two types of spikes: narrow, axonal spikes and broad spikes that are presumed to propagate into the dendrites [Bell et al., l997a]. Depression of the synapse between parallel fibers and the dendritic tree is observed if a broad spike occurs within a narrow window of time (up to 60 msec) following a parallel fiber spike. If the broad spike is absent, or occurs outside the time window, the parallel fiber spike appears to produce a non-associative enhancement of the synaptic efficacy [Bell et al., l997b]. The dynamics of this type of temporal learning rule lead to the storage of a negative image of sensory input and will be investigated in a companion paper [Roberts and Bell, 1999].

The purpose of the present article is to show that differential Hebbian learning naturally arises from the physiological learning rules of the type found by Markram et al. (1997b) and Zhang et al. (1998). The results show that these dynamics of synaptic plasticity are due to the dominantly *anti*symmetric structure of these temporal learning rules that generates opposite change in synaptic efficacy depending on whether the postsynaptic spike is in advance of, or follows, the presynaptic spike.

In the next section, we introduce the basic formalism to be used in our analysis, introduce the simulation method, and establish our the notation. The third section investigates the computational consequences of a specific temporal learning rule that has been measured in the neocortex [Markram et al., l997b]. The mathematical formalism is applied to the example of a repeated pairing of stimuli, and a simplified model of classical conditioning is used to demonstrate the resultant dynamics. We conclude with a brief discussion of the advantages and limitations of our approach, and some of the biological consequences of the results.

2. Mathematical Methods

The two mathematical approaches that will be used to study synaptic plasticity in the neocortex will be presented in this section. The first is analytic and uses a stochastic version of the synaptic response model [Gerstner and van Hemmen, l992] to reveal the origin of the differential learning rule. The second approach uses Monte Carlo simulation methods to illustrate the dynamics of synaptic change. The simulations complement the formal analysis when results are difficult to obtain analytically and serve to check the analytic results when the two methods overlap.

2.1 Stochastic Spike Response Models. Markram et al. (1997b) showed that synaptic plasticity in the neocortex is dependent on the relative timing of pre- and postsynaptic spikes. It is therefore important to calculate the probability of a postsynaptic spike as a function of time relative to the presynaptic spike. A spike occurs whenever the membrane potential of a neuron exceeds a threshold, $V(t) > \theta$. Because the membrane potential is influenced by many random processes beyond the control of the investigator, we assume that the membrane potential is a random variable with a normal (Gaussian) distribution function:

$$P(V - V_0, \sigma)dV = \frac{1}{\sigma\sqrt{2\pi}} \exp[-\frac{(V - V_0)^2}{2\sigma^2}]dV \quad (1)$$

where V_0 is the mean value of the membrane potential and σ^2 is the variance. The probability, $f_s(t)$, of a spike at time t is the probability that the membrane potential is greater than the threshold θ,

$$f_s(t) = f_s(V_0(t)) = \int_\theta^\infty P(V - V_0(t), \sigma)dV. \quad (2)$$

This is the defining expression for the *complementary error function* that is sigmoidal in form and has a value of 1/2 if $V_0(t) = \theta$. The firing rate of the model neuron is computed by dividing the probability by the refractory period. The absence of a membrane potential reset after each spike simplifies the analysis, and we have checked in the simulations that both model yield similar results.

Our interest is in the response of the neuron to a specific set of synaptic inputs that are either correlated in time with a known stimulus or in some way under the control of an experimental investigator. All other synaptic inputs are considered random and absorbed into a noise term. The contribution of excitatory postsynaptic potentials (epsps) is summed for all correlated synaptic inputs and weighted to compute the average membrane potential. The synaptic response function [Gerstner and van Hemmen, 1992] (or waveform) of each epsp (shown in Fig. 1B) is denoted by $E_i(t)$, where the index i identifies different synapses with decay times that range from 20 msec to 80 msec [Markram et al., 1997a]. These functions are normalized so that their integrals over time are unity, $\int_0^\infty E_i(t)\ dt = 1$, and when multiplied by a weighting factor, $w_i(t)$, yield the epsp.

In addition to the long term effects of pairing pre- and postsynaptic spikes, there may be a short term depression in these synapses that is dependent on the recent activity of each synapse [Abbott et al., 1997, Tsodyks and Markram, 1997, Zador and Dobrunz, 1997]. Although these short-term effects are not addressed in the following section, such depression can be represented by another weighting of the synaptic response function in the form of a signal function [Kosko, 1992], $S_i(t)$, that contains information about the effect of prior presynaptic activity on synapse i at time t. If one absorbs the background membrane potential into the threshold, the average membrane potential will only contain the contributions from temporally correlated synaptic inputs,

$$V_0(t) = \sum_i w_i(t) \int_0^\infty E_i(t') S_i(t-t') dt'. \qquad (3)$$

Short-term effects are important to the dynamics of biological neural networks and will be studied using this formalism in future investigations. In the following application, if the presynaptic spike frequency is low enough, the signal function can be represented by a sum of Dirac δ-functions [Dirac, 1958], $S_i(t) = \sum_m \delta(t-t_m)$, where t_m is a member of the set of spike arrival times at synapse i, and the sum is up to the longest time that yields a significant contribution from the synaptic response function to the average membrane potential in eq. (3).

The effects of a temporal learning rule on a longer time scale become evident by calculating the *average* change in synaptic weights after the pairing of pre- and postsynaptic activity. The change in the weight of a synapse that receives a presynaptic spike at time t, $\triangle w_n(t)$, is determined by the time, t_p, of a postsynaptic spike through a learning function, $L(t)$, describing the temporal learning rule, $\triangle w_n(t) = L(t_p - t)$ (such as shown in Fig. 1A). The *average* change in the synaptic weight, $< \triangle w_n(t) >$, is found by integrating over the probability of a postsynaptic spike at time t_p [Abbott and Blum, 1996],

$$< \triangle w_n(t) >= \int_{T_-}^{T_+} L(t_p - t) f_s(t_p) dt_p . \qquad (4)$$

The integration limits (T_-, T_+) are determined by the specific application. Typically, the learning function, $L(t)$ is negligible everywhere except for a small region near $t = 0$, so the integration limits may follow accordingly. It should be stressed here that the above expression is not deterministic in the sense that specific spikes at time t_p cause a change in the weights at time t. On the contrary, the expression simply provides a method to calculate the average change in the weights given the spike probability function during the time that the learning rule function $L(t_p - t)$ is non-zero. In the following application to the synapses of the neocortex, it will be shown how the form of the temporal learning rule on the time scale of milliseconds affects the functioning of neurons on the scale of seconds.

2.2 Monte Carlo Simulations. The simulations in the following section were designed to estimate changes in the activity patterns after specific training procedures and to display the results. The analytic results provide functional relationships among variables of the neural dynamics, but the simulations illustrate the synaptic change by applying individual postsynaptic spikes using the temporal learning rule during each training cycle. Over time, the effects of individual spikes average out and converge to the analytic results.

Custom simulation software has been written to carry out the simulations. The random number generator used a standard pseudo-random algorithm considered to be sufficient for simulations

of this scale [Binder, 1979]. The error function of eq.(2) was estimated using a linear segment in the neighborhood of $V_0(t) = \theta$ with a slope determined by the variance of the spike density function.

The sequence of operations that the computer program executed was as follows: First the modulation of the membrane potential due to the synaptic input throughout a training cycle was computed. Next, a postsynaptic spike was assigned to each time step in accordance with the spike probability function. Finally, the synaptic weights were updated by scanning the region around a specific temporally correlated synaptic input for a postsynaptic spike within the range of the temporal learning rule. If a spike was found, then the synaptic weight was depressed or enhanced by an amount dictated by the temporal learning rule (see Fig. 1). The sequence was then iterated until terminated by the investigator. Results were also obtained where the average membrane potential was reset after each spike.

3. Application: Neocortical temporal learning rule

The classical conditioning protocol consists of a repeated conditioned stimulus followed by the presentation of an unconditioned stimulus or reward. To formalize this protocol, the time component of the previous section will be split into two separate components, (x, t), where x is the time within each stimulus-reward cycle, and t represents the number of trials. This separation is appropriate because the measurable changes in behavior occur during the course of several training cycles (t), whereas the neuronal activity modulation that is responsible for synaptic change is greatest within each cycle (x).

In the analyses and simulations presented in this section, it is assumed that there is a series of sequentially delayed impulses arriving at separate synapses (Fig. 2) that are each temporally correlated with the conditioned stimulus. This assumption has been used previously to describe the neuronal models of classical conditioning [Buonomano and Mauk, 1994, Gluck and Thompson, 1990, Montague et al., 1996, Moore et al., 1989] and is referred to as a complete serial-compound stimulus [Sutton and Barto, 1990] or a spectral timing model [Grossberg and Schmajuk, 1989]. Under this *serial delay* assumption, the synaptic weights are labeled by their location in time following the onset of a stimulus, x, and change their values as several trials progress. Thus we denote the weights by $w(x, t)$. The underlying speculation in the serial delay assumption is that there is some mechanism that delays neuronal responses to a specific stimulus for a long enough time to link the stimulus to the time of the reward. To avoid short term depression effects, it is further assumed in this section that there is only one epsp from each synapse during the stimulus cycle.

In the next subsection we prove our central result; that the combination of an antisymmetric learning rule with a series of delayed synaptic inputs implies a differential Hebbian rule. Although the global learning dynamics of the model can be deduced directly from the differential Hebbian result, in the following subsection we use stability analysis to deduce the dynamics directly from the temporal learning rule $L(x, t)$. It will be shown that the neocortical synaptic learning rule results in travelling waves of neuronal activity during conditioning that propagate from the time of a later stimulus to the onset of an earlier, associated stimulus. The third subsection demonstrates the travelling wave in a computer simulation.

3.1 Differential Hebbian learning. The temporal learning rule that will be used is based on the results reported by Markram et al. (1997b) and Zhang et al. (1998). The data presented in first article do not fully characterize the details of the temporal learning rule, but give three points of the learning rule: enhancement of the synaptic efficacy if the epsp begins 10 msec before the postsynaptic spike, depression if the epsp begins 10 msec after the postsynaptic spike, and no change if the epsp and the postsynaptic spike were 100 msec apart. These data suggest that the model function be antisymmetric about the origin $(L(-x) = -L(x))$ and vanish for large positive and negative values of its argument. This model learning rule is further validated by the more complete results of the second article [Zhang et al., 1998] which reveal a learning rule that is nearly antisymmetric in time and vanishes for large $\pm x$.

The first derivative of the Gaussian distribution function fills the necessary requirements and has

the added advantage that the integral is normalized so as to separate the width, σ_L, from the learning rate, β. Thus, the learning rule is modeled as:

$$L(x) = -\beta P'(x, \sigma_L) = \frac{\beta x}{\sigma_L^3 \sqrt{2\pi}} \exp[-\frac{x^2}{2\sigma_L^2}]. \quad (1)$$

The extrema of $L(x)$ are at $x = \pm\sigma_L$. The value used in the following, $\sigma_L = 14$ msec, is consistent with the data reported in [Markram et al., 1997b].

We are now in the position to solve for the average change in synaptic efficacy, eq.(4), during each stimulus-reward cycle. The limits of integration will be taken to infinity because the value of $L(x)$ becomes indistinguishable from the noise for large $\pm x$ when β is assigned physiologically realistic values. In the continuous limit, the spike probability function is expanded about $z = x - x_p = 0$ to yield,

$$\begin{aligned} < \triangle w(x,t) > \; &\approx \; \int_{-\infty}^{\infty} L(z)[f_s(x,t) + z\frac{\partial}{\partial x} f_s(x,t) \\ &\quad + \frac{z^2}{2!}\frac{\partial^2}{\partial x^2} f_s(x,t)]dz \\ &= \; \sum_m L_m \frac{\partial^m}{\partial x^m} f_s(x,t), \end{aligned} \quad (2)$$

where the moments of the temporal learning rule are defined by

$$L_m = \frac{1}{m!}\int_{-\infty}^{\infty} z^m L(z) dz. \quad (3)$$

For an antisymmetric temporal learning rule such as that measured in [Markram et al., 1997b] and parametrized by eq.(1), $L_{2m} = 0$. The rule given in eq.(1) yields $L_1 = \beta$, and the higher order moments diminish rapidly.

The main result of this article is that the *average* change in synaptic efficacy is proportional to the *rate-of-change* of the postsynaptic spike probability,

$$< \triangle w(x,t) > \approx \beta \frac{\partial}{\partial x} f_s(x,t). \quad (4)$$

The reason for the importance of this result is that it represents the macroscopic results (on the time scale of several conditioning cycles) that follow from the microscopic temporal rule (on the time scale of the interspike spike interval). Again we must emphasize that the quantities of this derivation are averaged changes that depend on the functional form of probability functions. This result does not represent the precise synaptic weight changes the occur following each spike. That change is formalized by eq.(1), or an approximation thereof, and is sensitive only to postsynaptic events within about 40 msec of the presynsptic spike. However, over the course of several stimulus cycles, or averaged over several similar neural systems, the weight changes at each synapse tend to follow the derivative of the postsynaptic spike activity at the point in time when its epsp begins.

The dynamics of this learning rule follow directly form the average synaptic changes given by eq.(4). Suppose two stimuli that converge on the same cortical neuron are paired as follows: The first is distributed in a serial delay with plastic synaptic connections to the neuron that follow an (approximately) antisymmetric learning rule. The second stimulus arrives after the first, and during the delayed inputs. If the second stimulus is strong enough to increase the membrane potential above threshold, then the spike probability will be increased and have an increasing slope immediately prior to the second stimulus. The synapses with epsps that arrive during this period of increasing slope will, on the average, increase their efficacy and thus contribute to an increasing spike probability immediately prior to *their* arrival time. Eventually, after many pairings of the two stimuli, the increasing efficacy of later synapses contribute to the increasing efficacy of the earlier synapses until the efficacy of the first epsp following the first stimulus is saturated. The net effect is a change in latency of the neuron's response so that at the beginning of the pairing, it responded only to the second stimulus, but after many cycles the neuron responds with a high spike probability to the first stimulus.

As the cortical neuron learns to respond to the first stimulus, the latency decreases. In the components chosen for our representation of the stimulus-reward cycle, (x,t), the decreasing latency will appear as a *travelling wave* of spike activity from the time of the second stimulus to the first. In the next subsection we apply stability analysis to the averaged effects of the antisymmectric learning rule on synaptic weights to prove that the travelling wave exists and to compute the wave's velocity. These dynamics are then demon-

strated in the following subsection with computer simulations.

3.2 Instabilities of the temporal learning rule. As a pyramidal cell alters its response to a sensory stimulus under the influence of a temporal learning rule, the pattern of adaptive responses will be determined by the form of the rule. To characterize the form of adaptive responses to sensory input, we further analyze the equation for the average change in synaptic weights (eq.4) by investigating solutions of the pattern of weight values for the synapses.

Oscillatory instabilities often arise in complex systems as described by this model. Therefore, we seek solutions of the form

$$< w(x,t) > \propto e^{ikx} e^{\lambda t}, \qquad (5)$$

where λ is a growth factor parametrizing the emergence of a pattern during the course of many trials and k is the wave number [Murray, 1989]. The relation between the growth of oscillations and their wave number can be found by substituting this solution into the formal continuous limit ($\triangle t \rightarrow dt$, $\triangle w \rightarrow dw/dt$) of the weight evolution formula (eq.4). The validity of this limit will be tested in the simulations of the next subsection. It is reasonable to assume here that the neuron spends most of the time near its threshold so that a piecewise linear approximation of $f_s(x)$ is appropriate;

$$\frac{d}{dt} < w(x,t) > = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} dx' L(x'-x) \times \int_0^{\infty} dx'' < w(x'',t) > E(x'-x''). \quad (6)$$

Using the substitutions, $y = x' - x''$ and $z = x - x''$, the expression for the growth factor is found to be

$$\lambda = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} dz e^{-ikz} L(y-z) \int_0^{\infty} dy E(y). \quad (7)$$

The epsp waveform will be approximated with the function $E(y) = -2P'(y, \sigma_E)$ for $y > 0$. The integrals can be solved to yield

$$\lambda = \frac{\beta}{\sqrt{2\pi}\sigma} \{ \sigma_L \sigma_E k^2 e^{-\frac{1}{2}(\sigma_L^2 + \sigma_E^2)k^2} + i\sigma_L k \sqrt{\frac{2}{\pi}} (1 - r(\sigma_E k)) e^{-\frac{1}{2}\sigma_L^2 k^2} \}. \quad (8)$$

where,

$$r(\sigma_E k) = \sigma_E k e^{-\frac{1}{2}(\sigma_E k)^2} \int_0^{\sigma_E k} e^{\frac{1}{2}y^2} dy. \qquad (9)$$

The appearance of a complex growth factor implies the existence of traveling waves with a velocity, $c = -(\beta\sigma_L/\pi\sigma)(1 - r(\sigma_E k)) \exp[-(\sigma_E k)^2/2]$. The fact that the velocity of traveling wave solution is negative implies that any disturbance of the system, such as a perturbation of the membrane potential, will propagate by way of changes in synaptic weights to an earlier point in time. As will be demonstrated in the next subsection, this propagation of a signal is how the differential Hebbian rule links events that are separated in time and will be demonstrated in the simulations of the next section.

The real part of the growth factor in eq.(8) identifies the most unstable wave number that is likely to develop into oscillations during the course of several stimulus cycles. By analyzing the dispersion relations [Murray, 1989] of this system we find the unstable wave number to be $\hat{k} = \sqrt{2/(\sigma_L^2 + \sigma_E^2)}$. In the neocortex, oscillatory instabilities will manifest as high frequency bursts of action potentials by pyramidal neurons. To determine the frequency of these oscillations, we set the time course of the epsp to be consistent with physiological data found in Markram el al. (1997a) so that $\sigma_E = 7$ msec. A graph of the oscillation frequency as a function of the time course of the temporal learning rule is shown in Fig. 3. It is interesting to note that the range of frequencies includes the naturally occurring oscillations that are associated with synchrony among cortical neurons [Singer, 1993].

3.3 Simulating classical conditioning. These simulations demonstrate the change in the response of a neuron to a sensory stimulus during conditioning. The reward is modeled as a depolarizing increase in the membrane potential so that postsynaptic spikes affect the synaptic efficacy if there is presynaptic activity. The reward stimulus follows the stimulus by about 1 sec. During the course of conditioning, the response of the neuron to the conditioned stimulus is seen to increase. In addition, the time course of forgetting is explored when the reward is silenced.

The serial delay assumption is used in the following simulations to divide the time following

the conditioned stimulus into equal segments, $\triangle x$. The signal function (eq. 3) now takes the form, $S_n(x) = f_{pre}(n\triangle x)\delta(x - n\triangle x)$, where $f_{\rm pre}(n\triangle x)$ is the probability of a presynaptic spike arriving at time $x_n = n\triangle x$ following the conditioned stimulus. At each time step, the arriving signal is correlated with the stimulus, and all uncorrelated synaptic inputs are absorbed into the background noise. These expressions yield the discrete form of the average membrane potential,

$$V_0\ (x_n, t_m) = \sum_{i=1}^{2} \sum_{x_n - x_s \in I_E} w_i(x_s, t_m) E_i(x_n - x_s). \quad (10)$$

The sum over different synaptic response functions indexed by i in eq.(10) refer to the synapses with the same spike arrival time. We have obtained similar results when only one synapse is assigned to each time step. The second sum is over the range I_E where $E(x_n)$ is nonvanishing. The synaptic weights are restricted to the range $1 \leq w_i(x_n, t_m) \leq 60$, and each simulation begins with the initial conditions $w_i(x_n, 0) = 5$ for all x_n.

The presynaptic neurons are assumed to have a specified probability of firing so that the longer latency neurons diminish in their firing probability in keeping with expected physiological constraints. Thus, there is a probability profile that determines the probable activity at each synapse (Fig. 4A). Since a synapse must be contributing an epsp to the membrane potential in order to change its synaptic weight, the probability profile influences the rate of plasticity at each synapse. In fact, the left-moving velocity of the traveling waves, c, analyzed in the previous subsection will be diminished proportionally to the firing probability of the presynaptic neurons.

In the simulation, the membrane potentials and thresholds are normalized to a percentage of the maximum membrane potential, $V_{max} = \sum_i \sum_n w_{max} E_i(x_n) + W_R R_{max}$. We have used $R_{max} = 30$ as the maximum value of the contribution of the reward to the membrane potential, $W_R = 30$ as the weight of the reward input, and $w_{max} = 60$. The postsynaptic spike probability is calculated using a piece-wise linearized approximation of the complementary error function (eq. 2) as shown in Fig. 4B for the values used in the conditioned simulation where $\theta = 20\%$ and $\sigma = 0.08$.

For technical reasons, it was convenient to normalize the temporal learning rule as $L(x_n) = -10\beta\sigma_L P'(x_n, \sigma_L)$ and similarly the epsp waveform as $E(x_n) = -20\sigma_E P'(x_n, \sigma_E)$. In keeping with the available physiological data [Markram et al., 1997a, Markram et al., 1997b], we have set $\sigma_E = 7$ msec and $\sigma_L = 14$ msec. A refractory period of one time-step (7 msec) was included in the simulation, but did not noticeably influence the results. The learning rate was set at $\beta = 3$.

A typical conditioning simulation is shown in Fig. 5. The reward input is given between 840 - 910 msec following the beginning of the stimulus in the form of a sharp, triangular peak in membrane potential. The grey-scale background depicts the value of the membrane potential where the darker shade represents a higher potential. The number of spikes per stimulus cycle is given as a function of the number of trials to show the increase in the neuron's response to the stimulus during pairing with the reward input. The postsynaptic spike activity follows closely to the darker regions of the membrane potential.

An important feature of the learned response is that the beginning of the spike burst progresses earlier in time in each successive trial. This behavior is the characteristic large-scale result of the antisymmetric learning rule. In the first trial the neuron responds to the reward alone, but in time the neuron gives a long burst of action potentials in response to the conditioned stimulus. The neuron's response follows an S-shaped acquisition curve during the training session in agreement with previously reported results from whole-animal data [Klopf, 1988, Spence, 1956]. The two key aspects of the model for reproducing this acquisition curve are (1) the temporal learning rule [Markram et al., 1997b] and (2) the serial delay assumption. The variations of the exact shape of the probability profile (Fig. 4) do not change the basic form of acquisition curve, and the model is quite robust to changes in the threshold and the level of noise.

If the stimulus cycle is continued after the reward input is withheld, then the response of the neuron to the conditioned stimulus begins to fade. The time course of conditioning, followed by for-

getting, is shown in Fig. 6 for a typical simulation run. The beginning of the burst of action potentials is coincident with the early presynaptic inputs after about 75 trails and remains there throughout the forgetting phase. The reward input is withheld after the 100^{th} trial, and the decay of the neuron's response is slow while the length of the burst slowly shortens. Eventually the synaptic weights are reduced by the temporal learning rule to their lowest level and training to another reward input can commence as before.

4. Discussion

The analyses and simulations presented in this article show that the temporal learning rule reported by Markram et al. (1997b), in combination with a specific network architecture, leads to a change in neuronal responses that link initial (serial delayed) stimuli with a latter (unconditioned) stimuli. An essential aspect of the network is the serial delay assumption that distributes the input of the conditioned stimulus in a series of separate epsps that are precisely correlated in time with the stimulus. If a later input within the range of the serial delay series increases the neuron's activity, then the neuron learns to increases its activity in response to the conditioned stimulus through synaptic change. During the course of several trials, the beginning of the burst propagates foward in time from the later input until it is coincident with the onset of the conditioned stimulus.

The conclusions that follow from our results provide an indirect approach to test the serial delay assumption experimentally. One possible *in vivo* test would be to record from a region of the somatosensory cortex in which there are neurons with receptive fields to a specific somatosensory stimulus. Axon collaterals of responsive neurons will presumably make synaptic contact with dendrites of other pyramidal cells, though the connections may be too feeble to induce an action potential in the postsynaptic cells (see Fig. 3). If a neuron is found that does not respond to the same stimulus, then pairing the somatosensory stimulus with a delayed depolarization current would produce results similar to the simulations presented above if the serial delay assumption is valid.

The connection between the antisymmetric learning rule and differential Hebbian learning (Eq.4 suggests a synaptic mechanism for the classical conditioning models discussed in the Introduction. Since differential Hebbian learning is an approximation to the time-difference model of classical conditioning [Sutton and Barto, 1981], then it is likely that antisymmetric temporal learning rules give the brain the average behavior that is consistent with this model.

Two relevant aspects of cortical anatomy and physiology have not been taken into account in this study. First, inhibitory interneurons that are found in this part of the brain have been left out of our model primarily because the temporal learning rule of the associated synapses have not yet been characterized. The inclusion of these inhibitory inputs would not be expected to change the results much unless their temporal learning rule differs greatly. Judging by the parallel form of the temporal learning rule of inhibitory inputs suggested by studies in the electrolateral line lobe of mormyrid electric fish (C. Bell, personal communication), one expects the learning rule to be similar in the neocortex as well. In this case, plasticity of inhibitory inputs would add more of the details found in classical conditioning experiments carried out with whole animals.

The second physiological detail to be excluded in our model that may play an important role in the conclusions that can be drawn from the results is the short term depression found at these synapses [Abbott et al., 1997, Tsodyks and Markram, 1997]. The model studied in this article was designed to avoid influences of this effect because it is as yet unclear what influence the short term reduction of the synaptic efficacy has on the temporal learning rule. However, short-term depression is likely to increase when synaptic strength increases [Markram and Tsodyks, 1996]. Furthermore, it is known that short term depression of the form observed in the neocortex makes the postsynaptic response more sensitive to *changes* in the presynaptic spike frequency than to the frequency itself [Abbott et al., 1997, Tsodyks and Markram, 1997, Zador and Dobrunz, 1997]. If the sensitivity to the derivative of the presynaptic spike activity is raised, the behavior comes to resemble the full, i.e. pre- and postsynaptic,

differential Hebbian learning rule that has been studied in the past [Klopf, 1986, Kosko, 1986].

In the present study, the dynamics of synaptic change for the learning rule described in Markram et al. (1997b) are able to link any two excitatory events that can be traversed by a serial delay input. However, this model lacks a mechanism to encode the delay from the first event to the second; if a neuron responds to the second event, then following training, the neuron will begin its response immediately following the first event. The timing of conditioned responses may require the influence of another anatomical structure external to the neocortex. Studies of classical conditioning that involve lesions of the cerebellum suggest that this structure plays an important role in determining the timing of conditioned responses [Perrett et al., 1993]. In addition, the cerebellum appears to possess a very different type of temporal learning rule based on associative long-term depression that would lead to different dynamics of synaptic change. It is possible that the association between events is stored in the cerebral cortex, but the relative timing is encoded by the cerebellum.

Our analytical results provide a deeper understanding of previous theoretical studies on asymmetric temporal learning rules [Abbott and Blum, 1996, Gerstner et al., 1993]. These studies found that neural networks with a learning rule inspired by NMDA-mediated long-term potentiation are efficient for storing temporal sequences [Gerstner et al., 1993] and are predictive of input somewhat ahead of the input's timing in the original training sequence [Abbott and Blum, 1996]. The simplicity of the present model allows one to describe a sequence in its simplest form: two events separated in time. The prediction of future elements in a sequence arises from temporal instabilities in the dynamical equations of synaptic change (eq.6). Information about future events tends to propagate backwards in time so that synaptic efficacy is increased in anticipation of later stimuli.

In [Abbott and Blum, 1996], important factors for the shift in the coding vectors were related to the parameters of the learning rule's asymmetry and the width and overlap of each neuron's tuning curve. From the point of view of the present analysis, each neuron will learn to predict future events depending upon the distribution of incoming signals from other neurons, a distribution that is functionally related to the tuning curves and the overall input sequence. In addition, the present model suggests the time course of active forgetting that one would expect in these more complex network model. As demonstrated in Fig. 6, learning takes place much faster than the system forgets previously learned sequences.

The formalism presented here has been designed to include further details of synaptic dynamics as results become available. The stochastic methods that have been used show how dynamics that depend on the exact timing of events on the scale of milliseconds can average out to produce nontrivial behavior on the time scale of seconds. The antisymmetric temporal strucutre of the learning rule characterized in the neocortex [Markram et al., 1997b] yields a form of differential Hebbian learning that gives a mechanistic prediction for principles of classical conditioning.

Acknowledgements

I would like to thank Gin McCollum, Gerhard Magnus, and Curtis Bell for discussions and helpful suggestions on the manuscript. This research was supported in part by National Institutes of Health grant R01-DC02482 awarded to G. McCollum and National Science Foundation grant IBN 98-08887 awarded to PDR.

References

[Abbott and Blum, 1996] Abbott LF, Blum KI (1996) Functional significance of long-term potentiation for sequence learning and prediction. Cerebral Cortex 6:406–416.

[Abbott et al., 1997] Abbott LF, Varela JA, Sen K, Nelson SB (1997) Synaptic depression and cortical gain. Science 275:220–224.

[Bell et al., 1997a] Bell CC, Caputi A, Grant K (1997a) Physiology and plasticity of morphologically identified cells in the mormyrid electrosensory lobe. J. Neurosci. 17:6409–6423.

[Bell et al., 1997b] Bell CC, Han V, Sugawara Y, Grant K (1997b) Synaptic plasticity in a cerebellum-like structure depends on temporal order. Nature 387:278–281.

[Bi and ming Poo, 1998] Bi Q, ming Poo M (1998) Precise spike timing determines the direction and extent of synaptic modifications in cultured hippocampal neurons. J. Neurosci. 18:10464–10472.

[Binder, 1979] Binder K (1979). Introduction: Theory and "technical" aspects of Monte Carlo simulations. In:

K Binder, ed. Monte Carlo Methods in Statistical Physics, pp. 1–45, Berlin. Springer-Verlag.

[Buonomano and Mauk, l994] Buonomano DV, Mauk MD (l994) Neural network model of the cerebellum: Temporal discrimination and the timing of motor responses. Neural Comp. 6:38–55.

[Debanne et al., l994] Debanne D, Gähwiler BT, Thompson SH (l994) Asynchronous pre- and postsynaptic activity induces associative long-term depression in area CA1 of the rat hippocampus *in vitro*. Proc. Natl. Acad. Sci. USA 91:1148–1152.

[Dirac, 1958] Dirac PAM (1958). The Principles of Quantum Mechanics. Oxford University Press, Oxford.

[Gerstner et al., l993] Gerstner W, Ritz R, van Hemmen JL (l993) Why spikes? Hebbian learning and retrieval of time-resolved excitation patterns. Biol. Cybern. 69:503–515.

[Gerstner and van Hemmen, l992] Gerstner W, van Hemmen JL (l992) Associative memory in a network of 'spiking' neurons. Network 3:139–164.

[Gluck and Thompson, l990] Gluck MA, Thompson RF (l990). Adaptive signal processing in the cerebellum: Models of classical conditioning and vor adaptation. In: MA Gluck, DE Rumelhart, eds. Neuroscience and Connectionist Theory. Erlbaum.

[Grossberg and Schmajuk, 1989] Grossberg S, Schmajuk NA (1989) Neural dynamics of adaptive timing and temporal discrimination during associative learning. Neural Networks 2:79–102.

[Gustafsson et al., 1987] Gustafsson B, Wigström H, Abraham WC, Huang YY (1987) Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. J. Neurosci. 7:774–780.

[Hebb, 1949] Hebb DO (1949). The Organization of Behavior. Wiley, New York.

[Klopf, l986] Klopf AH (l986). A drive-reinforcement model of single neuron function: An alternative to the Hebbian neuronal model. In: JS Denker, ed. AIP Conference Proceedings 151:Neural Networks for Computing, pp. 265–270, New York. American Institute of Physics.

[Klopf, l988] Klopf AH (l988) A neuronal model of classical conditioning. Psychobiol. 16:85–125.

[Kosko, 1992] Kosko B (1992). Neural Systems and Fuzzy Systems. Prentice Hall, Englewood Cliffs.

[Kosko, l986] Kosko B (l986). Differential Hebbian learning. In: JS Denker, ed. AIP Conference Proceedings 151:Neural Networks for Computing, pp. 277–288, New York. American Institute of Physics.

[Levy and Steward, 1983] Levy WB, Steward O (1983) Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. Neurosci. 8:791–797.

[Markram et al., l997a] Markram H, Lübke J, Frotscher M, Roth A, Sakmann B (l997a) Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. J. Physiol. 500:409–440.

[Markram et al., l997b] Markram H, Lübke J, Frotscher M, Sakmann B (l997b) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science 275:213–215.

[Markram and Tsodyks, l996] Markram H, Tsodyks M (l996) Redistribution of synaptic efficacy between neocortical pyramidal neurons. Nature 382:807–810.

[Mauk, 1997] Mauk MD (1997) Roles of cerebellar cortex and nuclei in motor learning: contradictions or clues? Neuron 18:343–346.

[McKernan and Shinnick-Gallagher, 1997] McKernan MG, Shinnick-Gallagher P (1997) Fear conditioning induces a lasting potentiation of synaptic currents *in vitro*. Nature 390:607–611.

[Montague et al., l996] Montague PR, Dayan P, Sejnowski TJ (l996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J. Neurosci. 16:1936–1947.

[Moore et al., 1989] Moore W, Desmond JE, Berthier NE (1989) Adaptively timed conditioned responses and the cerebellum: A neural network approach. Biol. Cybern. 62:17–28.

[Murray, 1989] Murray JD (1989). Mathematical Biology. Springer-Verlag, Berlin.

[Pavlov, 1927] Pavlov IP (1927). Conditioned Reflexes. Oxford University Press, Oxford.

[Perrett et al., l993] Perrett SP, Ruiz BP, Mauk MD (l993) Cerebellar cortex lesions disrupt learning-dependent timing of conditioned eyelid responses. J. Neurosci. 13:1708–1718.

[Roberts and Bell, 1999] Roberts PD, Bell CC (1999). Computational consequences of temporally asymmetric learning rules: II. sensory image cancellation. (submitted).

[Rogan et al., 1997] Rogan MT, Stäubli UT, LeDoux JE (1997) Fear conditioning induces associative long-term potentiation in the amygdala. Nature 390:604–607.

[Singer, 1993] Singer W (1993) Synchronization of cortical activity and its putative role in information processing and learning. Ann. Rev. Physiol. 55:349–374.

[Spence, 1956] Spence KW (1956). Behavior Theory and Conditioning. Yale University Press, New Haven.

[Sutton and Barto, 1990] Sutton RS, Barto AG (1990). Time-derivative models of Pavlovian reinforcement. In: M Gabriel, J Moore, eds. Learning and Computational Neuroscience, pp. 497–537, Cambridge. The MIT Press.

[Sutton and Barto, l981] Sutton RS, Barto AG (l981) Toward a modern theory of adaptive networks: Expectation and prediction. Psychol. Rev. 88:135–170.

[Tsodyks and Markram, 1997] Tsodyks MV, Markram H (1997) The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc. Natl. Acad. Sci. USA 94:719–723.

[Zador and Dobrunz, 1997] Zador AM, Dobrunz LE (1997) Dynamic synapses in the cortex. Neuron 19:1–4.

[Zhang et al., 1998] Zhang LI, Tao HW, Holt CE, Harris WA, ming Poo M (1998) A critical window for cooperation and competition among developing retinotectal synapses. Nature 395:037.

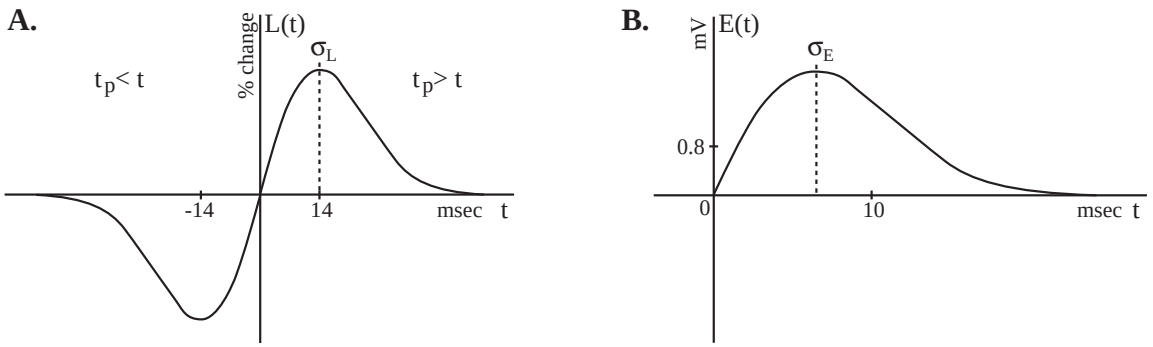

Fig. 1. Learning Rule. (A) The learning rule used in this study that is consistent with data presented in (Markram et al. 1997b). (B) The waveform of the response kernel similar to that found in (Markram et al. 1997a).

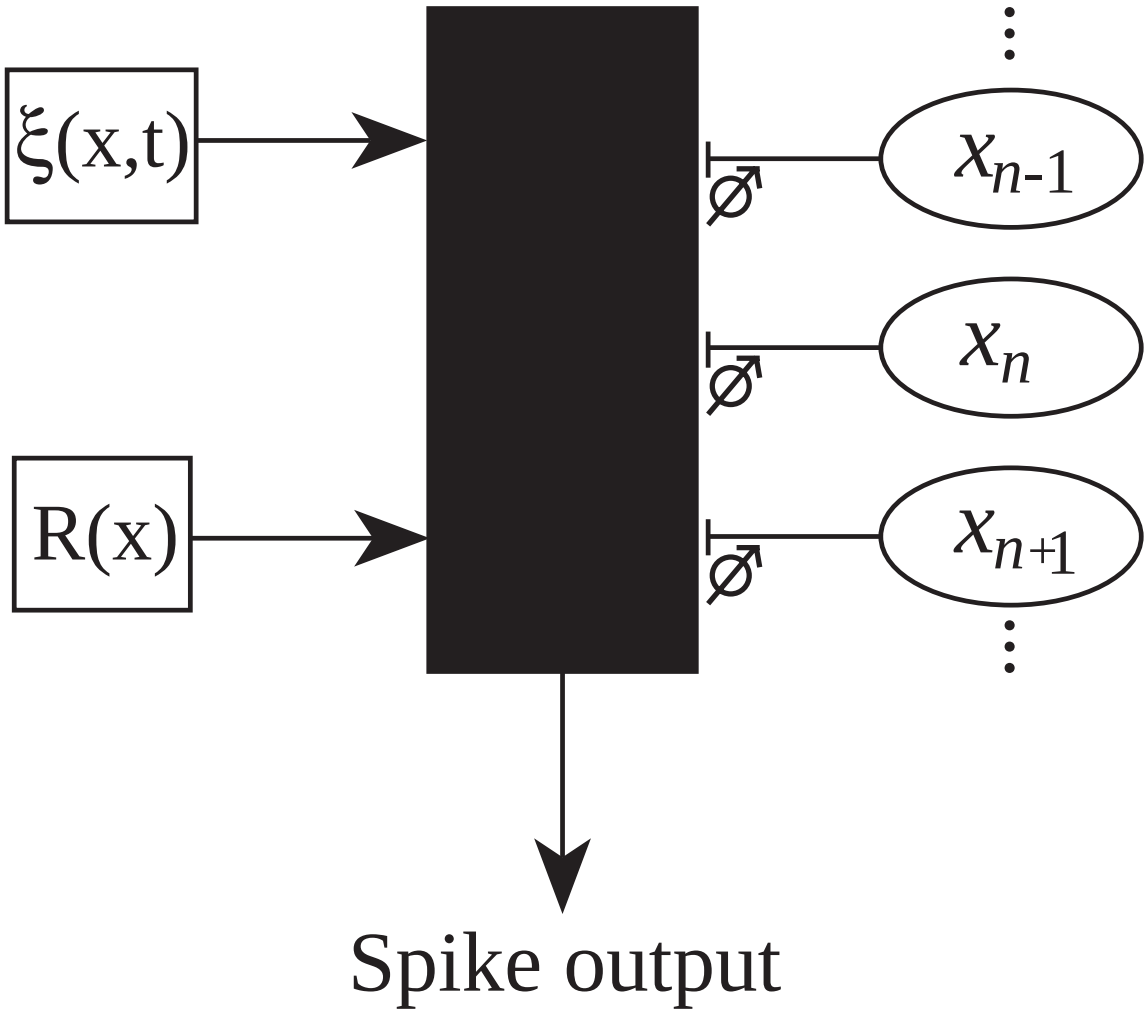

Fig. 2. The Model Neuron. A series of presynaptic neurons make excitatory synaptic contact with the model neuron so that the arrival time, x_n, of each epsp is consistently correlated with a conditioned stimulus during each cycle. t. These epsps are variable and change according to the learning rule described in the text. The model neuron is under the influence of noise ($\xi(x,t)$) including synaptic inputs that are not correlated with the conditioned stimulus, and internally induced variations in membrane potential. The unconditioned stimulus, or reward, $R(x)$, is introduced in the form of a small depolarizing input.

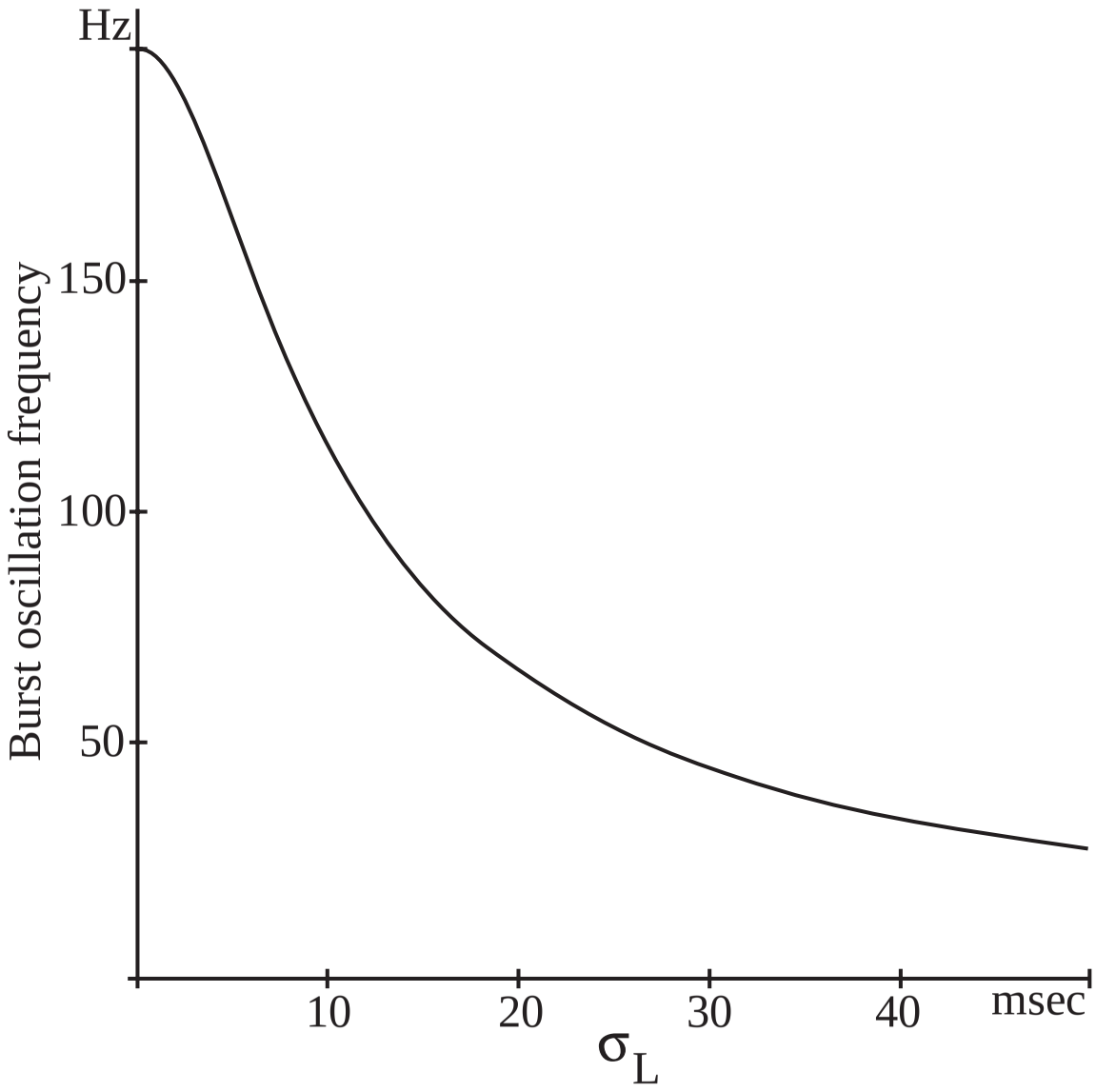

Fig. 3. Frequencies of instabilities. Instabilities that result from the temporal learning rule produce oscillatory responses of neurons to sensory stimulation. The frequencies of these oscillations are functionally dependent on a parameter, σ_L, of the temporal learning rule that determines the pre- and postsynaptic spike time differences that cause the greatest synaptic change.

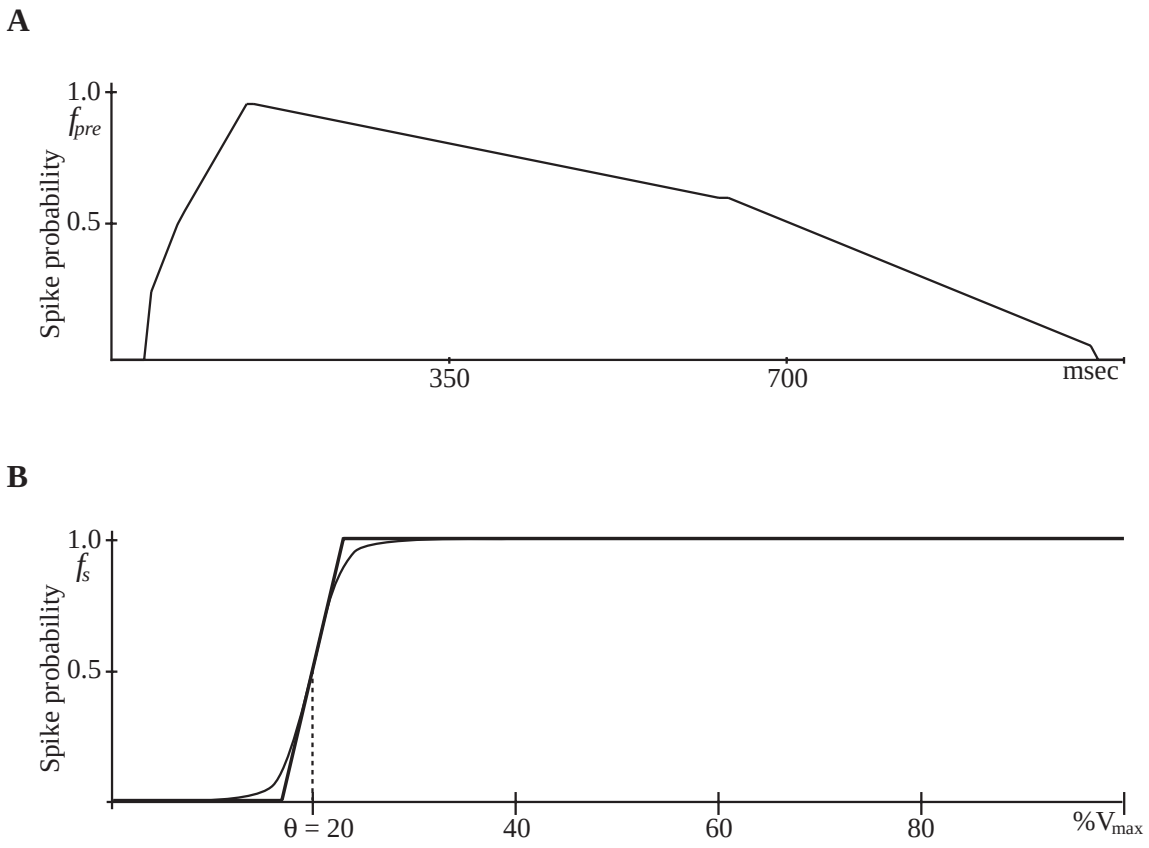

Fig. 4. Simulation functions. (A) Presynaptic spike probability profile as a function of arrival time at a synapse. (B) Postsynaptic spike probability approximated by a piece-wise linear function of the membrane potential. The potential is normalized as a percentage of the maximal calculated membrane potential.

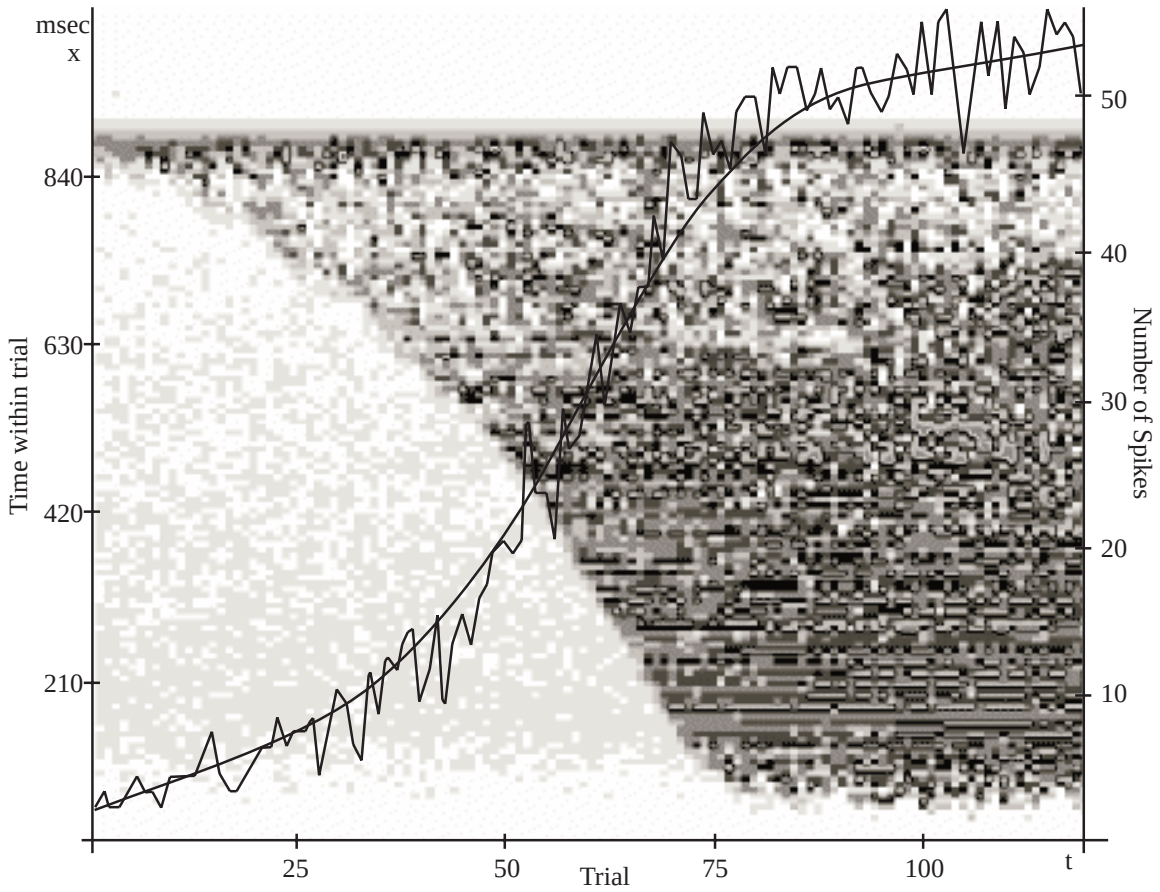

Fig. 5. Classical Conditioning. The trials run as a raster plot with the trial number across the bottom. The scale on the left gives the time following the stimulus onset and relates to the grey scale representation of the membrane potential in a typical simulation run. The darkness of the shade represents the magnitude of the membrane potential at each time step following the stimulus onset. The first trial begins with a reward input at $x_n = 850$ that briefly increases the membrane potential. As the trials progress, synaptic plasticity causes the front of the increase in membrane potential to travel forward in time until it reaches the onset of the stimulus input. This is the "travelling wave" discussed in the text. The S-shaped graph denotes the number of postsynaptic spikes during each stimulus and measures the response of the neuron to the stimulus cycle by the scale on the right. The smooth curve is the average spike number per stimulus cycle.

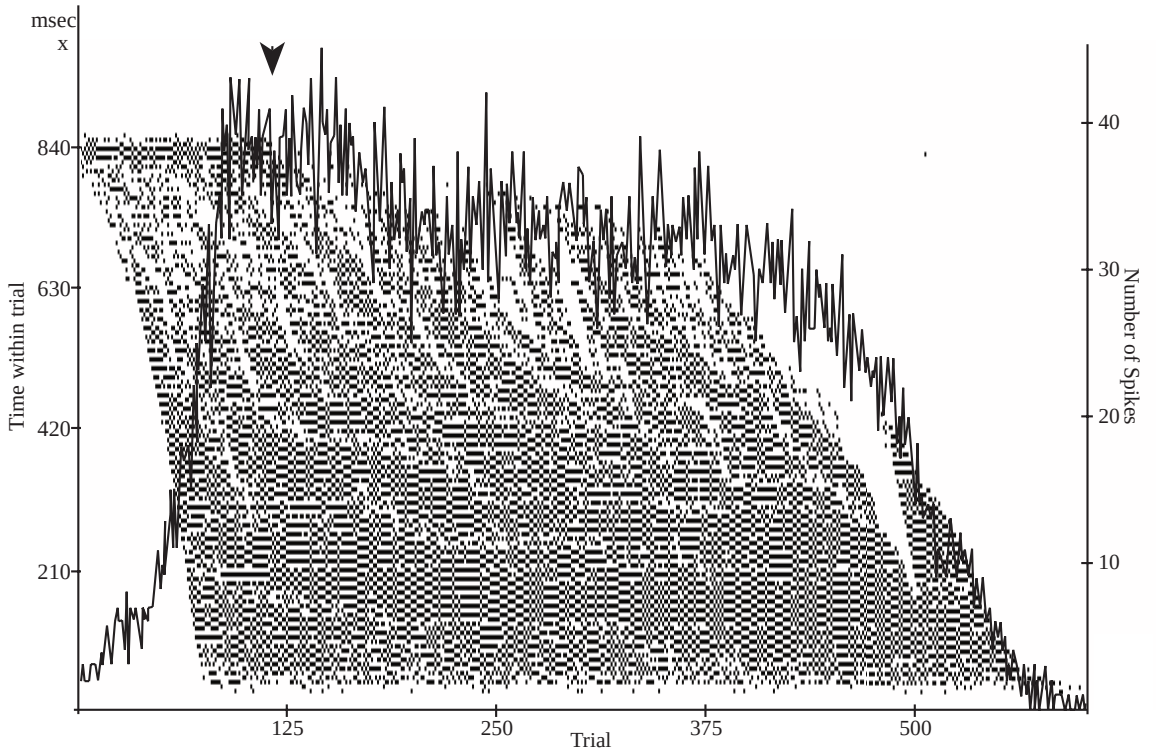

Fig. 6. Classical Forgetting. The number of spikes per stimulus cycle where the reward stimulus is presented through trial 100 (arrowhead) overlaid on a record of the spikes for a typical simulation run. The system retains the memory of the reward for many more trials than it took to reach maximal response to the conditioned stimulus. During the forgetting phase, the length of the burst decreases as the temporal learning rule reduces the weight of synapses with signals arriving at the end of the burst.