PHYSICAL REVIEW E VOLUME 62, NUMBER 3 SEPTEMBER 2000

Dynamics of temporal learning rules

Patrick D. Roberts
Neurological Sciences Institute, OHSU, 1120 Northwest 20th Avenue Portland, Oregon 97209

(Received 13 October 1999; revised manuscript received 13 March 2000)

The changes of synaptic strength are analyzed on two time scales: the fast local field dynamics, and the slow synaptic modification dynamics. The fast dynamics are determined by the synaptic strengths and background noise in the system. The slow dynamics are determined by the functional form of a *temporal learning rule*. Temporal learning rules are defined to be functions yielding state dependent changes in synaptic strengths depending on the timing of pre- and postsynaptic states in the network. The evolution of local field dynamics that result from various learning rules are analyzed for a stochastic, discrete time neural model with no relative refractory period that receives a series of delayed adaptive inputs. A fixed point is found in the learning dynamics, and conditions for two types of instabilities are analyzed. Four universality classes of dynamics are found that are independent of the details of the temporal learning rules. Examples are given of biological systems in which these temporal learning rules have been identified, and their functional consequences are discussed.

PACS number(s): 87.18.Sn, 87.19.La, 75.10.Nr

I. INTRODUCTION

In this paper we analyze the dynamics of learning in adaptive neural networks with generalized interactions between spin states (neurons). The interactions (synapses) considered here are dependent on the recent history of states, in contrast to the instantaneous state dependency of Hopfield networks [1,2]. In addition, the synapses are modified (learn) by a function of the time delay between the active states of the pre- and postsynaptic neurons. This generalizes Hopfield's implementation [1] of the Hebb rule [3] where the change of synaptic strength is a function of the averaged simultaneous activity of the pre- and postsynaptic neurons. These generalizations are shown to lead to static, oscillatory, and travelling wave solutions. Bifurcation parameters are found to be expressed in terms of the synaptic response functions and learning functions.

This work is part of an effort to generalize our understanding of learning dynamics to conditions that are important in biological neural networks [4–7]. The analysis and simulations presented here allow us to classify the types of dynamics that may arise when the interactions are not synchronous, but depend on recent states of the system. The stored memories of the system may be stabilized by including a temporal component to the learning rule that controls the changes of synapses. These types of learning rules are called *temporal learning rules* and have recently been characterized in biological systems [8–10].

The dynamics of adaptive neural networks can be separated into two time scales [11–13]: the fast activity of the neural states (response time), and the slow change in the strength of synapses between neurons (learning time). In the study of disordered systems, this separation of time scales is called partial annealing [12]. Here we follow the approach developed in [6] and apply separation of time scales to the analysis of learning in spiking neural networks.

The evolution of synaptic strength determines the dynamical classes. Under certain conditions a fixed point in the spike output probability function exists and the synaptic configuration is uniquely determined [14]. A Hopf bifurcation destabilizes the network's approach to this fixed point. The instability is due to the history dependent synapses, and conditions for stable learning will be derived in the following. Furthermore, traveling wave solutions are found for the equations describing synaptic modifications, and conditions for the appearance of these travelling wave solutions are identified. The learning dynamics show a universality that is independent of the exact form of the synaptic response function or temporal learning rule.

Although temporal learning rules explicitly describe the deterministic effect on the system by each pairing of states, the *average* changes in system states over time generate characteristic temporal patterns for each rule. Different temporal learning rules are then associated with identifiable dynamics. Four classes of dynamics will be treated in the following: stable approach to an equilibrium level of activity, oscillatory instabilities, and two traveling wave solutions. These dynamics fully describe the possible activity patterns that are generated by the connectivity explored in this study, and may be generalized to more complex systems.

The classes of dynamics follow from the formalization of phenomena found in biological neural tissue. However, the results uncover some interesting dynamics that are inherent in the fatigue of materials far from equilibrium whenever interactions change in a state-dependent manner. The adaptive dynamics are analyzed by using the simplest possible model that possesses the interesting dynamics; a set of temporally correlated states that interact with a stochastic ''neuron'' that is simultaneously being influenced by a time-dependent perturbation. As the perturbation is repeated, the strengths of the synapses change by rules that depend on the state of the whole system.

In the next section we present the network model and introduce our analytic approach. General temporal learning rules are formalized and the average changes of the synapses under temporal learning rule are established. The following section explores the state dynamics of the model neuron given a (nonbiological) rectangular response function. The

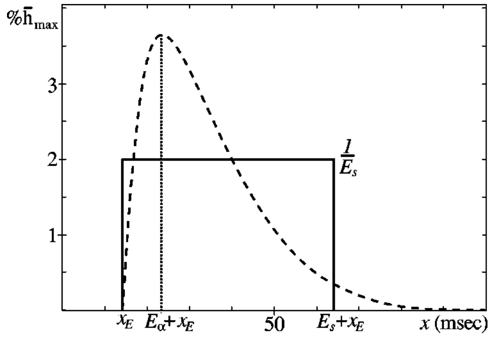

FIG. 1. *Response functions.* The functional form of the response functions are shown here. The square response function (solid) is used in the analysis of Sec. III, E_s parametrizes the width, X_s denotes the beginning of the nonzero segment, and the normalization constrains the height to $1/E_s$. The alpha-function response function (broken trace) is studied in Sec. IV, and has a maximum value at $E_\alpha + X_s$. The percentage of the maximum average local field value, $\%\overline{h}_{\max}$, is the value used in the simulations with $\Delta x = 1$ ms and the weights range from 0 to 1.

third section investigates the dynamics that arise from a more complicated, biologically realistic synaptic response function, and the theoretical significance of these results are presented in the final section.

II. THEORETICAL FRAMEWORK

To analyze the dynamics of temporal learning rules, we consider a two-state, stochastic neuron with (spin) states $s = 0$, 1. A series of I input sites are chronologically activated from $0 \rightarrow 1$ for one time step and returned to the resting state. This architecture is equivalent to that found in Fig. 1 of Kempter *et al.* [6]. However, in the present case each input site is activated at a different delay following a reference time, similar to the auditory scenario considered in [5]. The reference time represents the repeated presentation of a temporal pattern. Let $x = 0$ be the reference time, and all positive values of $x_i = i\Delta x$, $i = 1,2, \ldots, I$ represent the beginning of each response to the activation of the input series represented by the spike response function $\epsilon(x)$ (see Fig. 1). The spike response function is defined such that there is an increases probability of a postsynaptic spike following each input. This architecture is chosen because it is the simplest scenario that exhibits the dynamics characteristic of temporal learning rules. The model shows the dynamics for any learning system where adaptable inputs are repeatedly correlated with a nonadaptive temporal pattern. In addition, this architecture is relevant to certain biological neural systems [5,15,16], where the response function represents the postsynaptic potential due to synaptic input.

A. Fast neural dynamics

To investigate how the synaptic strengths of the input series change due to specific temporal learning rules, we wish to correlate the delayed series with a nonadaptive temporal pattern, $\xi(x)$. Each repetition of the paired input will be parametrized with the variable t that represents the evolution of the system under the influence of the learning rules. Thus, the neuron will be dependent on the coordinates (x,t), where the x-coordinate parametrizes fast neuronal (spin) dynamics, and the t-coordinate parametrizes slow learning (interaction) dynamics. We emphasize that x is not a spatial component, but the notation is used to suggest techniques from field theory.

We will compute the ensemble average change in synaptic strength following the pairing of inputs to the neuron. In each interval $[x_i, x_{i+1}]$, there is a finite probability of a *spike* $(s=1)$ that is functionally related to the average local field, $\overline{h}(x_i,t)$. Let $n(x_i)$ be the number of spikes between x_i and x_{i+1}. If β parametrizes the noise in the model neuron, the partition function given by

$$Z(\theta) = \prod_{i=1}^{I} \left(\sum_{n(x_i)} \exp[-\beta n(x_i)(\overline{h}(x_i,t) - \theta)] \right), \quad (2.1)$$

where θ is introduced as a spike threshold to regulate the average number of spikes, $\langle N(t) \rangle$, during each cycle.

If we choose the time steps of the x component to be as small as the absolute refractory period of the neuron, $r = \Delta x$, then only one spike can occupy each time-step, $n(x_i) = 0,1$. Thus

$$Z(\theta) = \prod_{i=1}^{I} [1 + \exp[-\beta(\overline{h}(x_i,t) - \theta)]]. \quad (2.2)$$

Using this partition function, we may compute the average number of spikes during each time step (see, for example, [17]):

$$\langle n(x_i,t) \rangle = -\frac{1}{\beta} \frac{\partial}{\partial h_i} \ln Z(\theta) = \frac{1}{1 + \exp[-\beta(\overline{h}(x_i,t) - \theta)]}, \quad (2.3)$$

where we have used the notation, $h_i = \overline{h}(x_i,t)$.

From the expression for the spike probability function, we see that given the noise parameter, β, and the spike threshold, θ, one may calculate the spike probability from the average local field, $\overline{h}(x_i,t)$. Each input of the delayed series contributes through the response function, $\epsilon(x_i)$, that represents the time course of the interaction. In the following, the response function will be normalized to unity, $\Sigma_i \epsilon(x_i) = 1$. When multiplied by a weight, $w(x_i,t)$, the product yields the contribution of a synapse that initiates its input at x_i. The total contribution from the series of adaptive inputs is the sum

$$h_\epsilon(x_j,t) = \sum_{i=1}^{N} w(x_i,t)\epsilon(x_j - x_i). \quad (2.4)$$

This quantity is added to the nonadaptive temporal pattern, $\xi(x_i)$, to yield the total average local field,

$$\overline{h}(x_i,t) = h_\epsilon(x_i,t) + \xi(x_i). \quad (2.5)$$

The weights are indexed by the arrive time of the presynaptic spike. Thus, $w(x_i,t)$ represents the weight that received a presynaptic spike at time x_i during cycle t, and the effect of that spike on the postsynaptic spike probability is represented by response function, $\epsilon(x_i)$.

B. Slow synaptic dynamics

We will now formalize the temporal learning rule in terms of changes in the synaptic strengths (synaptic weights), $w(x_i,t)$, over time *t*. Let s_i be the state of the input site that begins at time x_i. Let $L(x_j-x_i)$ be proportional to the change of the synaptic strength with input state $s(x_i)=1$ a postsynaptic spike at x_j. We restrict the domain of the associative rule to the *x*-component of our model (the fast time scale). Whenever possible, we will normalize the learning function such that $\Sigma_i L(x_i)=1$, and scale with a learning rate constant, λ.

Suppose that as *t* progresses the model neuron's spike output converges to a fixed point of the spike probability function denoted by $\hat{f}(x_i)$. The average change in the weights during each cycle will be dependent on the deviation of the spike probability function from $\hat{f}(x_i)$. This ensemble average change of the weights is then formalized by averaging the deviation from the final state over the learning rule:

$$\langle \Delta w(x_i,t)\rangle = \lambda \sum_j L(x_j-x_i)\big(\hat{f}(x_j)-\langle n^{out}(x_j,t)\rangle\big), \tag{2.6}$$

where the sum is over the nonvanishing extent of $L(x_i)$, and $\Delta w(x_i,t)=w(x_i,t)-w(x_i,t-1)$. The right hand side of this equation contains the information about the correlation between the pre- and postsynaptic spikes. However, this is implicit because our labeling of the weights. The input spike train at the synapse $w(x_i,t)$ is $n_i^{in}(x_j,t)=\delta_{ij}$ so that the input spike always occurs at the same time during the presentation of the nonadaptive temporal pattern.

Since the first term of Eq. (2.6) is independent of the state of the (postsynaptic) model neuron, it represents a nonassociative term. An important case is when the fixed point of the spike probability function is constant, $\hat{f}(x_i)=\hat{f}$. Then the average change in weights becomes

$$\langle \Delta w(x_i,t)\rangle = \lambda \hat{f} - \lambda \sum_j L(x_j-x_i)\langle n^{out}(x_j,t)\rangle. \tag{2.7}$$

This learning rule causes the system to approach an equilibrium of synaptic strengths that generates a negative image of the nonadaptive pattern, $\xi(x_i)$. The negative image is sculpted from the series of delayed inputs, $h_\epsilon(x_i)$, if certain constraints are imposed on the functional form of $L(x_i)$. We will investigate these constraints in the next two sections. Equation (2.7) is a special case of the learning rule studied in [6]. In the following we analyze the dynamics when the input spikes through the adaptive input are repeatedly correlated in time with a nonadaptive input.

Another important class of temporal learning rules is the case where associative enhancement and depression components cancel such that $\Sigma_i L(x_i)=0$. In this case, the average synaptic change simply becomes

$$\langle \Delta w(x_i,t)\rangle = -\lambda \sum_j L(x_j-x_i)\langle n^{out}(x_j,t)\rangle. \tag{2.8}$$

This learning rule cannot drive the system to a fixed point probability function that is constant in x_i.

We shall first explore a simple example of the dynamics associated with Eq. (2.6). Let $L(x_j-x_i)=\delta_{ji}$, the limiting case of a "temporal" learning rule where the synaptic change responds to postsynaptic activity at only one point in time. This equivalent to the learning rule used in Hopfield networks, except that here the synapses are history dependent making this a "spiking" model as opposed to a "rate" model. The average weight change is then

$$\langle \Delta w(x_i,t)\rangle = \lambda\big(\hat{f}(x_i)-\langle n^{out}(x_j,t)\rangle\big). \tag{2.9}$$

Let $\overline{h}^{\mathrm{eq}}(x_i)$ be the local field associated with the fixed point of the spike probability function, such that $\hat{f}(x_i) = f(\overline{h}^{\mathrm{eq}}(x_i))$. Expand $\langle n^{out}(x_j,t)\rangle$ about the deviation of the local field from the fixed point field, $\overline{h}(x_i,t)=\overline{h}^{\mathrm{eq}}(x_i)$, in the expression for the average synaptic change, Eq. (2.6). The lowest order term of the expansion yields

$$\langle \Delta w(x_i,t)\rangle = -\lambda Y(x_i)\big(\overline{h}(x_i,t)-\overline{h}^{\mathrm{eq}}(x_i)\big)+\cdots, \tag{2.10}$$

where $Y(x_i)=\beta \hat{f}(x_i)(1-\hat{f}(x_i))$. Defining a set of weights, $\{\hat{w}(x_i)\}$, such that the fixed point local field,

$$\overline{h}^{\mathrm{eq}}(x_i)=\sum_{j=1}^{N}\hat{w}(x_j)\epsilon(x_i-x_j)+\xi(x_i), \tag{2.11}$$

we have the expression for the average change in synaptic strengths,

$$\langle \Delta w(x_i,t)\rangle = -\lambda Y(x_i)\sum_{j=1}^{N}\big(\langle w(x_j,t)\rangle-\hat{w}(x_j)\big)\epsilon(x_i-x_j). \tag{2.12}$$

C. Stability analysis

The change of an individual weight in Eq. (2.12) depends on the strength of synapse that are in turn determined by other weights in the series of adaptive inputs. Thus, the nonlocality of the response function in the *x* component leads to instabilities that will drive the system away from the fixed point of the spike probability function. To investigate the conditions under which the instabilities arise, we substitute into Eq. (2.12) an oscillatory solution for the weight configuration in the *x* component. The *t* component is given an exponential decay parametrized by the constant τ. Stability will result if the decay constant is positive. The solution to be tested is

$$\langle w(x_i,t)\rangle = \hat{w}(x_i)+e^{ikx_i}e^{-t/\tau}, \tag{2.13}$$

where *k* is the wave number of instabilities that deviate from the fixed point of the weight configuration.

The computation is simplified if we approximate the sums with integrals, take the limit $\Delta x \to 0$, and $\Delta w \to (d/dt)w$. The problem now becomes a one-dimensional field theory in *x* that is evolving in time, *t*. Substituting Eq. (2.13) into this approximation of Eq. (2.12) yields

$$\frac{1}{\tau}=\lambda Y(x)\int \epsilon(y-x)e^{ik(y-x)}dy. \tag{2.14}$$

We will represent the response function by the so-called α-function with a variable decay constant: $\epsilon(x)=E^2e^{-Ex}$ for $x\geqslant 0$, otherwise $\epsilon(x)=0$ (shown in Fig. 1). The normalization is chosen so that $\int \epsilon(x)=1$. Using this functional form for the synaptic response function, Eq. (2.14) becomes

$$\frac{1}{\tau}=\lambda Y(x)E^2\frac{(E^2-k^2)+i2kE}{(E^2+k^2)^2}. \tag{2.15}$$

The system will spontaneously oscillate with a frequency of $1/k$ if the real part of $1/\tau$ is negative. Since $Y(x)$ is always greater than zero, $1/\tau$ is negative for values of k that are greater than E. Thus, the evolution of the weight configuration is unstable except in the limit of $E\rightarrow\infty$, i.e., $\epsilon(x)=\delta(x)$, the Dirac delta function. In addition, the complex term in Eq. (2.14) implies the existence of traveling wave solutions in the weight configuration so that the system never settles to a constant configuration.

This example is instructive for two reasons. First, it captures all of the dynamics that appear with temporal learning rules: a stable approach to the fixed point of the spike probability function in the limit, $E\rightarrow\infty$, and instabilities for all finite values of E. Second, it shows that instabilities develop if the temporal learning rule is not extended in time. The learning rules of the next two sections will generalize this case to study the dynamics when the rate of synaptic change is dependent on the activity at different times.

III. SQUARE RESPONSE FUNCTION

We now investigate the case in which both the learning rule and response functions are square waves. This case will show the effect of the learning rule's temporal dependence. This form of learning rule exhibits the universality classes containing the dynamics generated by the all temporal learning rules such as the biologically realistic α-function that will be treated in the next section. Let

$$\epsilon(x)=\begin{cases}1/E & \text{if} \quad x_E\leqslant x\leqslant x_E+E,\\ 0 & \text{otherwise},\end{cases} \tag{3.1}$$

where x_E is a real number that denotes the onset of the interaction (spike arrival), and E denotes the extent of the interaction (postsynaptic response to the spike) in time (see Fig. 1). This definition is constructed so that the area under $\epsilon(x)$ is unity. The learning function, $L(x)$, is defined similarly with E replaced by L. To simplify the calculation, we will choose the fixed point of the spike probability function to be a constant, $\hat{f}(x)=\hat{f}$. This choice of the fixed point does not detract from the generality of the results because we are primarily interested in instabilities of the equation of synaptic change that are dependent on the associative term.

A. Stability: Temporal pattern inverse

Proceeding as in the previous section, we now write Eq. (2.10) as

$$\langle \Delta w(x_i,t)\rangle=-\lambda Y\sum_j L(x_j-x_i)(\overline{h}(x_j,t)-\overline{h}^{\mathrm{eq}}(x_j)), \tag{3.2}$$

where Y is defined following Eq. (2.10). After substituting in the solution given in Eq. (2.13), and using the continuum approximation, we change variables to arrive at

$$\frac{1}{\tau}=\lambda Y\int\int L(y-z)\epsilon(y)e^{-ikz}dydz. \tag{3.3}$$

As before, the integrals are over the nonvanishing segments of $L(y-z)$ and $\epsilon(y)$. Carrying out the integration yields

$$\frac{1}{\tau}=\lambda Y\frac{4}{k^2EL}\sin\left(\frac{kE}{2}\right)\sin\left(\frac{kL}{2}\right)e^{ik(X_E-X_L+L/2-E/2)}. \tag{3.4}$$

The real part of this expression is positive for all k only if $E=L$ and $X_E=X_L$.

Thus, stability of the temporal learning rule is assured only when there is an exact match between the learning function and the response function. Under these conditions, the synaptic strengths relax until their inputs exactly cancel the nonadaptive temporal pattern. When the learning parameters are changed, a Hopf bifurcation is crossed and oscillations appear.

B. Instabilities: Oscillations and traveling waves

We may generalize our representation of the associative component of the learning rule by separating the learning function into a segment that reduces the weights, $L^-(x)\geqslant 0$ for all x, and a segment that increases the weights, $L^+(x)\geqslant 0$ for all x. The average change in weights now becomes

$$\langle \Delta w(x_i,t)\rangle=\lambda\hat{f}-\lambda^+\sum_j L^+(x-x_i)\langle n^{out}(x_j,t)\rangle$$
$$-\lambda^-\sum_j L^-(x-x_i)\langle n^{out}(x_j,t)\rangle, \tag{3.5}$$

where $L^{\pm}(x)=\mp 1/L^{\pm}$ for $X_{\pm}\leqslant x\leqslant X_{\pm}+L^{\pm}$. Without loss of generality, we may scale the x-component such that $E=1$, and set the beginning of each synaptic response so that $X_E=0$. When we carry out the same calculation as above, we find that the real part of the decay constant is

$$\mathrm{Re}(\tau)=\frac{4}{Yk^2}\left[\frac{\lambda^-}{L^-}\cos\left(\frac{k}{2}(2X_-+L^--1)\right)\sin\left(\frac{k}{2}L^-\right)\right.$$
$$\left.-\frac{\lambda^+}{L^+}\cos\left(\frac{k}{2}(2X_++L^+-1)\right)\sin\left(\frac{k}{2}L^+\right)\right]\sin\left(\frac{k}{2}\right). \tag{3.6}$$

The stable approach to equilibrium of the synaptic strengths under the influence of the learning rule requires that this expression is positive for all k.

Two important categories of learning rules depend on the region of the learning rule that is coincident with the interaction. Hebbian learning is defined as a positive synaptic change during the interaction, and anti-Hebbian learning is defined as a negative synaptic change. In the first case, the associative increase of the synaptic strength is coincident with the response function so that $L^+=1$ and $X_+=0$. Sub-

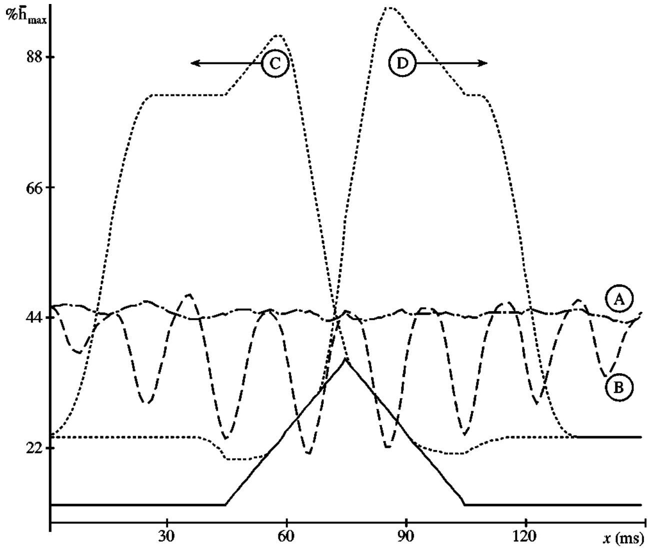

FIG. 2. *Instabilities.* The average local field as generated by four different temporal learning rules in a simulation. In all cases, $L^- = L^+ = E_s$, and the nonadaptive input (solid line) was introduced at $t=0$. (A) Only associative enhancement ($\lambda^+=0$), $\lambda\hat{f}>0$, $X_-=0$, and $t=200$. (B) Same as (A) with $X_+=-E_s/2$, and $t=1000$. (C) $\lambda\hat{f}=0$, $\lambda^+=\lambda^->0$, $X_+=0$, $X_-=E_s$, and $t=20$. (D) Same as (C) with $X_-=-E_s$ and $t=20$.

stituting these values into Eq. (3.6) we find that $\mathrm{Re}(\tau)<0$ for some value of k regardless of our choice of the remaining parameters. Thus, Hebbian learning causes perturbations from the fixed point to grow with time so that nonadaptive temporal patterns are magnified by the adaptive inputs.

The case of anti-Hebbian learning is more stable, as can be seen when we set $L^-=1$ and $X_-=0$. It is possible for $\mathrm{Re}(\tau)\geqslant 0$ for all k if $\lambda^->\lambda^+$ for a wide range of values of L^+ and X_+. In fact, if $L^+=1$, then

$$\mathrm{Re}(\tau)=\frac{4}{Yk^2}[\lambda^- - \lambda^+\cos(kX_+)]\sin^2\left(\frac{k}{2}\right). \tag{3.7}$$

This implies stability regardless of the time delay between the beginning of the synaptic response and the onset of the weight increase in the learning function. Due to the strong stability that results from anti-Hebbian learning, when attempting to store a target temporal pattern it is perhaps better to base learning algorithm on anti-Hebbian learning rather than the traditional implementations of the Hebbian rule [3]. However, the anti-Hebbian system will not ''complete'' an input pattern, but the stored pattern is recovered through a reversal of the output when there is no input pattern.

The types of the instabilities reflected by the average local field, $\bar{h}(x,t)$, are shown in Fig. 2. This figure was generated by a computer simulation of the model's dynamics. At the beginning of the simulation, a nonadaptive input, $\xi(x)$, is delivered for the time steps, $45<x_n<105$, shown by the solid line. When there is an exact match between the response function and learning function [$\epsilon(x)=L(x)$ for all x], and the nonassociative learning rate, $\lambda\hat{f}>0$, then the adaptive input weights adjust to cancel the nonadaptive input pattern, $\xi(x)$. The result is trace (A) in the figure. However, when the learning rule is symmetric about the origin, then the oscillations shown by trace (B) develop. The left- and right-moving traveling waves (C and D) are generated by learning rules where $\lambda^-=\lambda^+$ and $\lambda\hat{f}>0$. Since these are the only instabilities that arise in one-dimensional systems, we have a full classification of the dynamics generated by temporal learning rules.

IV. BIOLOGICAL RESPONSE FUNCTION

In this section we will investigate the dynamics that result when the learning rule is represented by α functions. This functional dependence of the learning rule is consistent with the known biological mechanisms thought to be responsible for long term changes in synaptic strength [18]. Let $\epsilon(x)$ be defined as in Sec. II, and define $L(x)$ also as an α functions with the parameter L replacing E. Using these functions in the dynamical equation of synaptic change, Eq. (3.3) becomes

$$\frac{1}{\tau}=\lambda Y E^2 L^2\int_0^y dz\int_0^\infty dy\,y(y-z)e^{-L(y-z)}e^{-Ey}e^{ikz}. \tag{4.1}$$

The integrals can be evaluated to yield

$$\frac{1}{\tau}=\lambda Y\frac{E^2L^2}{(ik+L)^2}\left[-\frac{2(ik+L)}{(L+E)^3}+\frac{1}{(ik-E)^2}-\frac{1}{(L+E)^2}\right]. \tag{4.2}$$

The conditions for stable solutions can be made more apparent by rescaling the x component in units of $E=1$. After some algebra, one finds that the real part of $1/\tau$ is negative if, for some choice of k,

$$\begin{aligned}0>{}&(1-L)k^6+(3-L+2L^2-2L^3)k^4-(2-3L+3L^2+L^3\\&+L^4-L^5)k^2+(L+L^2+L^4).\end{aligned} \tag{4.3}$$

This condition is satisfied, and thus the rule is unstable, for all $L>E$. If $L\leqslant E$, a narrow band of values exist where there are no solutions for the wave number such that Eq. (4.3) is negative (shown in Fig. 3). Thus, the α-function is slightly more forgiving than the square response function. The other results from the analysis for the square response function still hold, but with the caveat that there is more tolerance for parameter deviations from an exact match between the response function, $\epsilon(x)$, and the learning function, $L(x)$, to insure stability.

V. DISCUSSION

The primary goal of this project is to understand how the dynamics of biological neural networks complicate known dynamics of adaptive neural networks. We have shown that given a series of delayed inputs, temporal learning rules lead to one of four dynamical types: (1) stable approach to an equilibrium activity pattern; (2) oscillatory instabilities with a frequency determined by learning rule parameters; (3) traveling waves propagating in the $+x$ direction; (4) traveling waves propagating in the $-x$ direction. As long as the time scale is such that the duration of the synaptic response is not

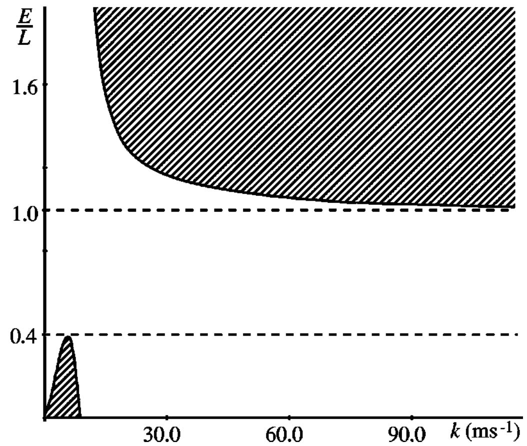

FIG. 3. *Range of stability for learning rules based on the α function.* The shaded regions are solutions for k where $1/\tau < 0$. Stability occurs only for values of L where the ratio E/L is between 0.4 and 1.

negligible, instabilities arise if enhancement of the synapses is associated with the postsynaptic response. Thus, there is no Hebbian learning in the traditional sense, only anti-Hebbian (activity pattern 1) and differential Hebbian learning (activity patterns 2 and 3).

The neural architecture investigated here is far simpler than that found in most neural networks (biological and artificial). Two important elements that are certain to have an important influence on the dynamics resulting from particular learning rules are recurrent connections and inhibition. This study must therefore be considered as a first-order analysis of the dynamics of temporal learning rules in biological systems. However, the simplification presented here preserves what are expected to be the major components of the dynamics, even if modified by recurrent connections and inhibition.

The first of these solutions has been observed in the electrosensory system of mormyrid electric fish, and there is reason to believe that the same principle underlies sensory processing in several other systems [19], such as the mammalian auditory system and the cerebellum [20]. Instabilities would disrupt the sensory images in these systems [14]. Thus, the fact that there is a region of stability in Fig. 3 is advantageous for these systems because a precise match between the associative depression of the learning rule and the response function is not required for a stable final image.

However, instabilities can have an important function in some biological systems. Learning rules leading to the traveling wave solution in the $-x$ direction have also been observed, and may serve an important function because this arrangement would help the brain to associate events that are separate in time. If a sensory input to a neuron is distributed in time through a series on adaptable synapses, and a later input is paired with the initial stimulus, then the activity initiated by the later input will propagate to the beginning of the distributed input. Thus the traveling wave solution acts to link the two events in time [7]. That this type of learning rule is found in cortical brain structures that are important in the association of different events is suggestive. One can expect that the form of the temporal learning rule plays a critical role in determining the function of each biological neural network.

Generalizations of the methods presented here will be of interest to other fields besides neuroscience. The recent interest in disordered systems out of equilibrium provides a rich background for the separation of the time scales into different variables. It would be interesting to compute variables such as energy and state correlations in larger networks with the learning rules and interactions studied here. The results of such a computation may provide insight into the aging of quenched materials under externally applied stress.

ACKNOWLEDGMENTS

The author would like to thank Gin McCollum, and Curt Bell for discussions and helpful suggestions on the manuscript. This research was supported in part by National Science Foundation Grant No. IBN-9808887.

[1] J. H. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **79**, 2554 (1982).
[2] J. H. Hopfield, Proc. Natl. Acad. Sci. U.S.A. **81**, 3088 (1984).
[3] D. O. Hebb, *The Organization of Behavior* (Wiley, New York, 1949).
[4] W. Gerstner, R. Ritz, and J. L. van Hemmen, Biol. Cybern. **69**, 503 (1993).
[5] W. Gerstner, R. Kempter, J. L. van Hemmen, and H. Wagner, Nature (London) **383**, 76 (1996).
[6] R. Kempter, W. Gerstner, and J. L. van Hemmen, Phys. Rev. E **59**, 4498 (1999).
[7] P. D. Roberts, J. Comput. Neurosci. **7**, 235 (1999).
[8] C. C. Bell, V. Han, Y. Sugawara, and K. Grant, Nature (London) **387**, 278 (1997).
[9] H. Markram, J. Lübke, M. Frotscher, and B. Sakmann, Science **275**, 213 (1997).
[10] L. I. Zhang, H. W. Tao, C. E. Holt, W. A. Harris, and M.-M. Poo, Nature (London) **395**, 037 (1998).
[11] D. Caroppo and S. Stramaglia, Phys. Lett. A **246**, 55 (1998).
[12] V. Dotsenko, S. Franz, and M. Mézard, J. Phys. A **27**, 2351 (1994).
[13] R. W. Penney, A. C. C. Coolen, and D. Sherrington, J. Phys. A **26**, 3681 (1993).
[14] P. D. Roberts and C. C. Bell, J. Comput. Neurosci. **9**, 67 (2000).
[15] C. C. Bell, K. Grant, and J. Serrier, J. Neurophysiol. **68**, 843 (1992).
[16] R. S. Sutton and A. G. Barto, in *Learning and Computational Neuroscience*, edited by M. Gabriel and J. Moore (The MIT Press, Cambridge, MA, 1990), pp. 497–537.
[17] R. P. Feynman, *Statistical Mechanics* (Benjamin, Reading, MA, 1972).
[18] D. Debanne, B. T. Gähwiler, and S. H. Thompson, Proc. Natl. Acad. Sci. U.S.A. **91**, 1148 (1994).
[19] C. C. Bell, D. Bodznick, J. Montgomery, and J. Bastian, Brain Behav. Evol. **50**, 17 (l997), (Suppl. 1).
[20] M. D. Mauk, Neuron **18**, 343 (1997).