Biol. Cybern. 64, 231–241 (1991)

Biological Cybernetics

Variable threshold as a model for selective attention, (de)sensitization, and anesthesia in associative neural networks

L. Wang and J. Ross

Department of Chemistry, Stanford University, Stanford, CA 94305, USA

Received January 31, 1990/Accepted in revised form August 27, 1990

Abstract. We study the influence of a variable neuronal threshold on fixed points and convergence rates of an associative neural network in the presence of noise. We allow a random distribution in the activity levels of the patterns stored, and a modification to the standard Hebbian learning rule is proposed for this purpose. There is a threshold at which the retrieval ability, including the average final overlap and the convergence rate, is optimized for patterns with a particular activity level at a given noise level. This type of selective attention to one class of patterns with a certain activity level may be obtained at the cost of reducing the retrieval ability of the network for patterns with different activity levels. The effects of a constant threshold independent of noise, time, and pattern are discussed. For high-(low-) activity patterns, the average final overlap is shown to be increased at high noise levels and decreased at low noise levels by a negative (positive) constant threshold, whereas a positive (negative) threshold always reduces the final average overlap. When the magnitude of the constant threshold exceeds a critical value, there is no retrieval. Rates of convergence towards the stored pattern with negative (positive) thresholds are greater than those with positive (negative) thresholds. These results are related to (de)sensitization and anesthesia. For certain threshold values and patterns with certain activity levels, hysteresis appears in the plot of the average final overlap versus the noise level, even for first order interactions. We make the analogy between the *pattern-dependent* neuronal threshold proposed in the present paper and the "*task-related*" modulation in neuronal excitability determined by cognitive factors, such as the attentional state of a higher animal. A constant threshold is associated with overall changes in neuronal excitability caused, e.g., by various drugs and physical injuries. Neurophysiological evidence of a dynamically variable neuronal threshold, such as accommodation and potentiation, is presented.

1 Introduction

Since the pioneering work of Grossberg (1970); Little (1974); Kohonen (1988); and Hopfield (1982, 1984), there have been many research efforts on improvements and generalizations of associative neural network models (for recent reviews, see e.g., Sompolinsky 1988; Domany 1988; Crick 1989). In conventional associative neural network models, the threshold for a neuron to fire an action potential is fixed (usually assumed to be zero), and stored patterns are created randomly. As a result, there are equal amounts of "on" bits and "off" bits in each pattern stored. This type of pattern is called unbiased and has an activity level, which is defined as the fraction of "on" bits, $r = 50\%$ (or 0.5). Amit et al. (1987a) have recently discussed neural networks with low levels of activity ($r < 0.5$). In their analysis stored patterns are biased and the system is said to be "magnetized" in the language of spin glasses. They have studied the so-called *constraint dynamics* in which there is a non-zero threshold for each neuron so that the system evolves only in certain regions of the phase space with pre-specified activity level. In the unbiased case, this imposed threshold vanishes. The work of Amit et al. (1987a) has been generalized by Gutfreund (1988) to a layered neural network model with hierarchically correlated patterns (Parga and Virasoro 1986; Cortes et al. 1987). Krogh and Hertz (1988) have presented a mean-field analysis on hierarchical neural networks with a variable threshold at a finite temperature, that is, in the presence of noise. They have shown that by varying the threshold it is possible to focus on a certain level in the hierarchy and increase the capacity slightly. A mean-field study has been carried out for an asynchronous neural network model with an adaptive threshold and biased patterns for the noise-free case (Buhmann et al. 1989). All the work cited concentrates on the storage capacity and the effect of a finite magnetization (bias in stored patterns); however, issues such as the detailed analysis on the dependences of the dynamics (the convergence rate) and the fixed points on noise and threshold of the network, as well as the neuro-physiological aspects related to a variable neuronal threshold, have not been discussed.

In the present paper we consider theoretically the effects of a variable threshold on convergence rates and fixed points of an associative neural network model in the presence of noise. We allow a random distribution in the activity levels of the patterns stored. Dynamical formulation with arbitrary threshold values is presented in Sect. 2, where a modification to the standard Hebbian learning rule is proposed for storage of patterns with an activity distribution. For patterns with a given activity level, we find, in Sect. 3, an optimal threshold as a function of the noise level and the time, at which the retrieval ability, e.g., the average final overlap and the convergence rate, of the network for these patterns are the greatest. While the neuronal threshold is set to its optimal value corresponding to patterns with one activity level, we show that performances of the network with respect to patterns with different activity levels are significantly reduced. Hence the network can achieve selective attention by choosing threshold values. The effects of a constant (noise-, time-, and pattern-independent) threshold are discussed in Sect. 4. For high- (low-) activity patterns, the average final overlap is shown to be increased at high noise levels and decreased at low noise levels by a negative (positive) threshold, whereas a positive (negative) threshold always reduced the final average overlap. When the magnitude of the threshold exceeds the certain critical value, there is no retrieval. These results are related to sensitization, desensitization, and anesthesia. For certain threshold values and patterns with certain activity levels, hysteresis appears in the plot of the average final overlap versus the noise level, even for first order interactions. Rates of convergence towards a stored pattern are studied as function of threshold in Sect. 5. Effects of a variable threshold on the memory capacity are studied in Sect. 6. In Sect. 7, we discuss neurophysiological evidence of a variable neuronal threshold.

2 Dynamical formulation

Consider a network of McCulloch-Pitts (1943) neurons connected via the Hebbian learning rule (Hebb 1949; Hopfield 1982). Suppose that the threshold for the i-th neuron is θ_i, which may depend on a number of variables such as time, noise and activity levels; applying an external field $-\theta_i$ to the i-th neuron is equivalent to shifting its threshold to θ_i. For simplicity, we consider a synchronous updating algorithm (Little 1974; Kinzel 1985; Keeler et al. 1989; Wang and Ross 1990 a, b); the dynamics of the system is

$$S_i(t+1) = \mathrm{sign}[h_i(t) - \theta_i]\,, \tag{1}$$

where $\mathrm{sign}(x) = -1$ for negative x and $\mathrm{sign}(x) = +1$ for positive x, and

$$h_i(t) = \sum_{j=1}^{N} T_{ij} S_j(t) + \eta_i\,, \tag{2a}$$

is the total input signal for the i-th neuron. We have included in (2a) a random Gaussian noise η_i with a mean of zero in order to take into account the presence of noise (Abelles 1982). In the presence of noise or a non-zero temperature, the dynamics is no longer deterministic. The discussions presented below are therefore probabilistic, i.e., we treat average behaviors in many independent trials (ensembles). The synaptic efficacy T_{ij} is specified below.

An interesting scaling property results from the simple fact that $\mathrm{sign}(x) = \mathrm{sign}(x/c)$, for any positive c. In the case where the synaptic efficacy undergoes a change from T_{ij} to $c_i T_{ij}$ for any arbitrary $c_i > 0$, this synaptic change can be rescaled into the threshold and the noise as follows

$$\begin{aligned} S_i(t+1) &= \mathrm{sign}\left[\sum_{j=1}^{N} c_i T_{ij} S_j(t) + \eta_i - \theta_i\right] \\ &= \mathrm{sign}\left[\sum_{j=1}^{N} T_{ij} S_j(t) + \frac{\eta_i}{c_i} - \frac{\theta_i}{c_i}\right]. \end{aligned} \tag{2b}$$

In conventional associative neural network models, each bit in the stored pattern is assigned "+1" or "−1" with equal probability. Thus the activity level of the μ-th pattern r_μ, defined as N_+^μ/N where N_+^μ is the number of "+1"s in the μ-th pattern, is 50%. In other words, $r_\mu = 0.5$ for $\mu = 1, 2, \ldots, p$, where p is the number of stored patterns. The deviations of the activity levels from 0.5 are on the order of $1/\sqrt{N}$, which is negligible for large networks considered here. In the present paper, we consider a more general case in which we allow a random distribution of activity levels in the stored patterns, i.e., $0 \le r_\mu \le 1$ for $\mu = 1, 2, \ldots, p$. A procedure to generate a collection of patterns with a distribution in their activity levels may be described as follows. First pick an activity level r_1 from a random number generator, which produces a random activity level. Assign each of the N bits of the pattern "+1" with a probability r_1 and "−1" with a probability $1 - r_1$. We thus have produced a random pattern $\mathbf{S}^1$ for this selection of activity level. Next pick another activity r_2 from the same random number generator and produce another pattern $\mathbf{S}^2$. We pick a total of p values of activity level to produce p patterns. In order to store patterns with an activity distribution, the original Hopfield prescription of the Hebbian learning rule (Hebb 1949; Hopfield 1982) has to be modified. We introduce the form, the motivation for which will be discussed shortly,

$$T_{ij} = \frac{1}{N}\left(\sum_{\mu=1}^{p} S_i^\mu S_j^\mu\right) - \frac{4(p-1)(\delta^2 + \Delta^2)}{N}\,, \tag{3a}$$

where S_i^μ is the μ-th stored pattern,

$$\delta^2 \equiv \langle r_\mu^2 \rangle_R - (\langle r_\mu \rangle_R)^2 \quad \text{for} \quad \mu = 1, 2, \ldots, p \tag{3b}$$

is the standard deviation of the activity distribution, and

$$\Delta \equiv \langle r_\mu \rangle_R - 0.5 \quad \text{for} \quad \mu = 1, 2, \ldots, p \tag{3c}$$

represents the bias of the average activity level; $\langle\rangle_R$ is an *average*[1] over the activity distribution. The quantities δ and Δ are the same for every pattern $\mu = 1, 2, \ldots, p$. In the absence of the activity distribution and average bias, i.e., $\delta = \Delta = 0$, (3) reduces to the standard Hebbian rule (Hebb 1949; Hopfield 1982).

Similar to our previous work (Wang and Ross 1990a, b), we consider the pattern $\mathbf{S}^1$ say, which is in the vicinity of the initial state of the network. Explicitly, we let the initial conditions of the network be

$$m^1(0) = \max\{m^\mu(0) | \mu = 1, 2, \ldots, p\}, \tag{4}$$

where $\max\{\cdots\}$ is the maximum in the data set given in the curly parentheses and

$$m^\mu(t) = \frac{1}{N} \mathbf{S}^\mu \cdot \mathbf{S}(t) \tag{5}$$

is the overlap between the state of the system at time t and the μ-th stored pattern. Note that $m^1(0)$ does not need to be 1; thus the initial condition of the network is not necessarily in complete overlap with pattern $\mathbf{S}^1$. We have defined the term *vicinity* by a scalar product given by (5) (the overlap); but one can also define vicinity through a norm or Hamming distance. These definitions are all equivalent (see e.g., Wang and Ross 1990a).

We separate the first term in the total signal (2) into two parts (Kinzel 1985):

$$\begin{aligned} h_i(t) = m^1(t) S_i^1 + \frac{1}{N} \sum_{j=1}^{N} \sum_{\mu=2}^{p} \Big[S_i^\mu S_j^\mu \\ - 4(\delta^2 + \Delta^2) \Big] S_j(t) + \eta_i ; \end{aligned} \tag{6}$$

the first term is proportional to the overlap of the system with pattern $\mathbf{S}^1$. The second part, the "crosstalk" (Keeler et al. 1989) term, is interference from patterns $\mathbf{S}^2, \mathbf{S}^3, \ldots, \mathbf{S}^p$. In a fully connected network such as the one discussed here, a detailed analysis of effects of the interference on network performance, especially the dynamics, is quite complicated (see e.g., Amit et al. 1985, 1987b). For a qualitative discussion, we adapt the approximation in which the correlation between the state of the system $\mathbf{S}(t)$ and patterns $\mathbf{S}^2, \mathbf{S}^3, \ldots, \mathbf{S}^p$ is neglected (see e.g., Kinzel 1985). It has been shown (Derrida et al. 1987) that this approximation becomes exact if enough synapses are disconnected randomly (dilution) in a large network. In this approximation, the crosstalk becomes a sum of $N(p-1)$ independent random variables with *zero averages*, which is proven as follows.

We recall that there are two random processes in generating the stored patterns, i.e., *(i)* select an activity level r for each pattern and *(ii)* select a random pattern according to this activity level. Thus there are two corresponding average processes. The first corresponds to random process *(i)* and is an average over the activity level distribution, i.e., $\langle\rangle_R$, which is used in (3b) and (3c). We denote the second average corresponding to random process *(ii)* by $\langle\rangle_b$, which stands for an average over many possible choices of the μ-th stored pattern according to a given activity level r_μ. Therefore the total average over the randomness in the stored patterns is $\langle\cdots\rangle \equiv \langle\langle\cdots\rangle_b\rangle_R$. Since all neurons are equivalent on average, for all $i = 1, 2, \ldots, N$, we have $\langle S_i^\mu \rangle_b = \langle (1/N) \Sigma_{i=1}^{N} S_i^\mu \rangle_b = 2r_\mu - 1$. For instance, at 50% activity level, or $r_\mu = 0.5$, the mean of the state variable for each neuron is $\langle S_i^\mu \rangle = 2r_\mu - 1 = 0$. Hence

$$\langle S_i^\mu S_j^\mu \rangle_b = \langle S_i^\mu \rangle_b \langle S_j^\mu \rangle_b = (2r_\mu - 1)^2 . \tag{7a}$$

The total average of this term is:

$$\langle S_i^\mu S_j^\mu \rangle \equiv \langle\langle S_i^\mu S_j^\mu \rangle_b \rangle_R = \langle (2r_\mu - 1)^2 \rangle_R = 4(\delta^2 + \Delta^2) . \tag{7b}$$

The last equality in (7b) is a direct result from (3b) and (3c). Therefore the averages of all interference terms in (6) vanish

$$\begin{aligned} &\langle [S_i^\mu S_j^\mu - 4(\delta^2 + \Delta^2)] S_j(t) \rangle \\ &\quad = \langle S_i^\mu S_j^\mu - 4(\delta^2 + \Delta^2) \rangle \langle S_j(t) \rangle = 0 . \end{aligned} \tag{7c}$$

Additionally, the standard deviation of each interference term is

$$\begin{aligned} &\langle [S_i^\mu S_j^\mu - 4(\delta^2 + \Delta^2)]^2 [S_j(t)]^2 \rangle \\ &\quad = \langle [S_i^\mu S_j^\mu]^2 \rangle - \langle S_i^\mu S_j^\mu \rangle^2 = 1 - 16(\delta^2 + \Delta^2)^2 . \end{aligned} \tag{7d}$$

Thus, in this approximation, the distribution of the interference becomes a time- and neuron-independent Gaussian with a deviation

$$\sigma_{ct} = \sqrt{\frac{(p-1)[1 - 16(\delta^2 + \Delta^2)^2]}{N}} \tag{7e}$$

in the large N limit (Kinzel 1985; Noest 1988), according to the central limit theorem.

We now discuss the motivation for modifying the standard version of the Hebbian learning rule, which is a prescription for determining synaptic efficacy T_{ij} in order to stabilize a given set of patterns with no activity distribution (Hebb 1949; Hopfield 1982). The modified learning rule given by (3) guarantees the stability (in an average sense, if there is noise) of a stored pattern in the presence of an activity distribution. Without losing generality, let us consider the stability of pattern $\mathbf{S}^1$. Suppose at a certain time the state of the network is in pattern $\mathbf{S}^1$, i.e., $\mathbf{S}(t) = \mathbf{S}^1$ at some time t, the "local field" for the i-neuron is obtained by letting $\mathbf{S}(t) = \mathbf{S}^1$ in (6). In the light of (7c), which results from the modified

---

[1] Mathematically, the average of a quantity Q over a random distribution $P(Q)$ is calculated by integrating this quantity with the distribution function over all possible values, i.e., $\langle Q \rangle = \int Q' P(Q')\, \mathrm{d}Q'$, where $P(Q')\, \mathrm{d}Q'$ is the probability that Q is between values Q' and $Q' + \mathrm{d}Q'$. Practically, e.g., in a computer simulation, the average of a given random variable is simply the mean of the values that this random variable assumes in many independent trials, provided that the number of trials is large. Suppose there are two random variables, Q_1 and Q_2, and $P(Q_1, Q_2)$ is their joint distribution function, i.e., $P(Q_1', Q_2')\, \mathrm{d}Q_1'\, \mathrm{d}Q_2'$ is the probability that Q_1 is between values Q_1' and $Q_1' + \mathrm{d}Q_1'$, and *at the same time*, Q_2 is between values Q_2' and $Q_2' + \mathrm{d}Q_2'$. Thus, $\langle Q_1 Q_2 \rangle = \int Q_1' Q_2' P(Q_1', Q_2')\, \mathrm{d}Q_1'\, \mathrm{d}Q_2'$. Random variables Q_1 and Q_2 are said to be uncorrelated if and only if $P(Q_1, Q_2) = P(Q_1) P(Q_2)$, so that $\langle Q_1 Q_2 \rangle = \int Q_1' Q_2' P(Q_1') P(Q_2')\, \mathrm{d}Q_1'\, \mathrm{d}Q_2' = \langle Q_1 \rangle \langle Q_2 \rangle$

learning rule given by (3), (6) indicates that the average "local field" is aligned with the "spin" when $\mathbf{S}(t) = \mathbf{S}^1$, i.e., $\langle \mathbf{h}(t) \rangle = \mathbf{S}^1$. This means that for any threshold $|\theta_i| < 1$, $\mathbf{S}^1$ is an *attractor* in the system, according to the update algorithm given by (1) (note that $\mathbf{S}^1$ *would be a repellor* if *we had* $\langle \mathbf{h}(t) \rangle = -\mathbf{S}^1$ instead). Stability of other stored patterns, i.e., $\mathbf{S}^2, \mathbf{S}^3, \ldots, \mathbf{S}^p$, can be demonstrated in the same fashion. Hence we have proven that the modified Hebbian learning rule given by (3) stabalize a given set of patterns with an activity distribution. Effects of a variable threshold on the memory capacity, i.e., maximum number of patterns that the network can store, will be discussed in Sect. 6.

We continue our derivation of the dynamical equation obeyed by the network. In the approximation that the interference terms in the double sum of (6) are independent random variables, we have shown that resultant random variable is Gaussian distributed and has a mean of zero. We recall that the external noise η_i is also a random variable with a Gaussian distribution and mean of zero. We can therefore combine the interference with the external noise to get a total noise of the same nature, i.e., with a Gaussian distribution and mean of zero. By substituting (1) and (6) into (5) (with $\mu = 1$), we obtain

$$m^1(t+1) = \frac{1}{N} \sum_{i=1}^{N} S_i \operatorname{sign}[S_i m^1(t) - \theta_i + \eta'] , \tag{8}$$

where η' is the combination of the internal noise and the interference with a deviation (Keeler et al., 1989)

$$\sigma = \sqrt{\sigma_{ct}^2 + \sigma_0^2} . \tag{9}$$

From here on, we omit superscript "1" and assume that the state of the neural network is converging to pattern $\mathbf{S}^1$, unless otherwise specified. Hence, under a mean-field approximation that the average can be taken inside the sign function, i.e., $\langle \operatorname{sign}(x) \rangle \approx \operatorname{sign}(\langle x \rangle)$, the average overlap can be written as follows (detailed proof is given in the Appendix)

$$\langle m(t+1) \rangle = \frac{1}{N} \sum_{i=1}^{N} \{1 - 2\psi[\langle m(t) \rangle - \theta_i S_i]\} , \tag{10}$$

with

$$\psi(y) = \frac{1}{\sqrt{2\pi}} \int_{y/\sigma}^{+\infty} \mathrm{e}^{-x^2/2} \, \mathrm{d}x . \tag{11}$$

In (11) and onwards, the statistical average $\langle \rangle$ includes not only an average over the randomness in the stored patterns, but also an average over the external Gaussian noise (η_i).

For simplicity, we now assume that all neurons in the network have the same threshold value, i.e., $\theta_i = \theta$, for all i; however, θ may depend on variables such as the time, the noise level, and the activity level. Suppose that there are N_+ "+1" and $(N - N_+)$ "−1" in $\mathbf{S}^1$. Separating the summation in (10) into two groups, with $S_i = +1$ and $S_i = -1$, respectively, we obtain the following dynamic equation for a neural network with a variable threshold θ

$$\begin{aligned} \langle m(t+1) \rangle = r\{1 - 2\psi[\langle m(t) \rangle - \theta]\} \\ + (1 - r)\{1 - 2\psi[\langle m(t) \rangle + \theta]\} , \end{aligned} \tag{12}$$

where $r \equiv N_+/N$ is the activity level of $\mathbf{S}^1$ (we have omitted the subscript "1") and $0 \le r \le 1$.

In the following sections, we study the consequences of a variable threshold by analyzing (12) in detail. Although we have assumed in (12) that every neuron has the same threshold, (12) is still quite general since the threshold θ can be an independent variable, or it can depend, in any arbitrary fashion, on other variables such as the time, the noise level, and the activity level. In the present paper we are concerned with only two special cases. In Sect. 3, we seek a special dependence of the threshold on the time, the noise, and the activity level so that the retrieval ability of the network is optimized with respect to a particular class of patterns. In Sect. 4, we evaluate effects of a constant (time-, noise-, activity-independent) threshold.

3 Optimal threshold: selective attention

The average overlap between the state of the network at time $t + 1$ given by (12) is a function of, among other variables, the threshold value θ. We now find an optimal threshold which maximizes the average overlap $\langle m(t+1) \rangle$ at each time step. Notice that

$$\begin{aligned} \frac{\partial \langle m(t+1) \rangle}{\partial \theta} = \frac{2}{\sqrt{2\pi}\,\sigma} \{ -r\, \mathrm{e}^{-[\langle m(t) \rangle - \theta]^2/2\sigma^2} \\ + (1 - r)\, \mathrm{e}^{-[\langle m(t) \rangle + \theta]^2/2\sigma^2} \} \gtreqqless 0, \quad \text{if} \end{aligned} \tag{13}$$

$$\theta \gtreqqless \theta_{\max} \equiv \frac{\sigma^2}{2\langle m(t) \rangle} \ln\left(\frac{1}{r} - 1\right) . \tag{14}$$

Therefore $\langle m(t+1) \rangle$ reaches the maximum for a given set of r (or N_+), σ, and t, when $\theta = \theta_{\max}$. In other words, the average overlap given by (12) is the largest at any given time if we choose $\theta = \theta_{\max}$. The final average overlap $\langle m(\infty) \rangle$, which is the fixed point of (12), as well as the convergence rate towards $\langle m(\infty) \rangle$, for which more discussions will be presented in Sect. 5, is therefore also the greatest. The limit $t = \infty$ always exists in the present system since the network has a Lyapunov (energy) function and always reaches thermal equilibrium (see e.g., Hopfield 1982).

We than concluded that the retrieval ability of the neural network is maximized for patterns with activity level r by choosing an optimal threshold $\theta = \theta_{\max}$ given by (14).

Several analytic conclusions concerning the fixed points of (12) can be drawn for the case where $\theta = \theta_{\max}$. First of all, we have

$$\begin{aligned} \frac{\partial y_1[\langle m(\infty) \rangle]}{\partial \langle m(\infty) \rangle} = \frac{2}{\sqrt{2\pi}\,\sigma} \{ r\mathrm{e}^{-[\langle m(\infty) \rangle - \theta_{\max})^2/2\sigma^2} \\ + (1 - r)\mathrm{e}^{-[\langle m(\infty) \rangle + \theta_{\max}]^2 2\sigma^2} \} > 0 , \end{aligned} \tag{15}$$

where

$$y_1[\langle m(\infty)\rangle] \equiv r\{1 - 2\psi[\langle m(\infty)\rangle - \theta]\} + (1-r)\{1 - 2\psi[\langle m(\infty)\rangle + \theta]\} \qquad (16)$$

is the right hand side of (12) with $\theta = \theta_{\max}$ and $t = \infty$. Equation (15) means that $y_1[\langle m(\infty)\rangle]$ increases monotonously as $\langle m(\infty)\rangle$ increases. Hence if we plot $y_1[\langle m(\infty)\rangle]$ and $y_2 \equiv \langle m(\infty)\rangle$ against $\langle m(\infty)\rangle$, the latter being a diagonal straight line, there exists only one intersecting point, which is the fixed point $\langle m(\infty)\rangle$ of (12), for a given set of r (or N_+) and σ: There can be no hysteresis in the plot of the final average overlap $\langle m(\infty)\rangle$ versus the noise level σ if $\theta = \theta_{\max}$, see Fig. 1.

Secondly, we show that

$$\frac{\partial y_1[\langle m(\infty)\rangle]}{\partial \sigma} = -\frac{2\langle m(\infty)\rangle}{\sqrt{2\pi}\,\sigma^2}\{re^{-[\langle m(\infty)\rangle - \theta_{\max}]^2/2\sigma^2} + (1-r)e^{-[\langle m(\infty)\rangle + \theta_{\max}]^2/2\sigma^2}\} < 0\,. \qquad (17)$$

Hence $y_1(\theta = \theta_{\max})$ decreases as σ increases: for a given r (or N_+), the curve $y_1(\theta = \theta_{\max})$ plotted against $\langle m(\infty)\rangle$ gets flatter and flatter as σ increases. This curve becomes a straight line, i.e., $y_1(\theta = \theta_{\max}, t = \infty) = |1 - 2r|$, for $\sigma = +\infty$. Hence, when intersected with $y_2 \equiv \langle m(\infty)\rangle$, the fixed point of (12) at the high noise level limit is

$$\langle m(t = \infty)\rangle = |1 - 2r|, \quad \text{for} \quad \sigma = +\infty\,. \qquad (18)$$

After solving (12) numerically in the case where $\theta = \theta_{\max}$, we plot in Fig. 1 the final average overlap $\langle m(\infty)\rangle$ versus the noise level σ for various choices of activity level r. The general behavior of these curves agrees with our analytic arguments presented above: At any non-zero noise level and for a pattern with an arbitrary activity level $r \neq 0.5$, the average final overlap with threshold set to $\theta_{\max}$ is always greater than the conventional result where a zero threshold is chosen [compare curves (a), (b) with curve (c)]. This improvement on the retrieval ability of the network is more pronounced at high noise levels. As the noise level increases, the average final overlap saturates at the value given by (18). There is no hysteresis in this plot.

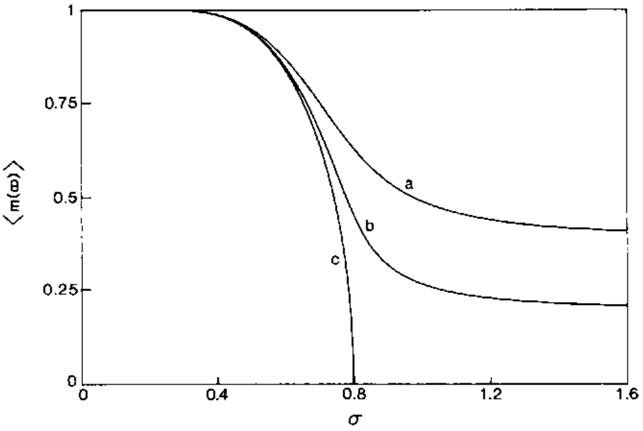

Fig. 1. Plot of average final overlap $\langle m(\infty)\rangle$ given by (12) vs. the standard deviation of the Gaussian noise σ for patterns with various activity levels: *(a)* $r = 0.7$ *(b)* $r = 0.6$ *(c)* $r = 0.5$. The threshold is set to $\theta = \theta_{\max}(r)$ given by (14). Curve *(c)* is the conventional plot, since when $r = 0.5$, $\theta = \theta_{\max} = 0$

As indicated by (14), there are only two possibilities that the conventional zero threshold can be the optimal threshold ($\theta_{\max} = 0$). The first is in the absence of noise, i.e., $\sigma = 0$, where all curves in Fig. 1 meet at the point $\langle m(\infty)\rangle = 1$. The second possibility is when $r = 0.5$ (or $N_+ = N/2$), which represents an unbiased pattern (curve (c) in Fig. 1).

The self-consistent solutions of (12) and (14) show that the optimal threshold $\theta_{\max}$ is negative (positive) for $r > 0.5$ ($r < 0.5$) and decreases (increases) monotonically as the noise level σ increases.

An interesting observation can be made through (18). The more biased (r closer to 1 or 0) the activity is, the larger the effect of letting $\theta = \theta_{\max}$. For instance, in Fig. 1, the activity level in curve (a), where $r = 0.7$, is closer to 1 than the activity level in curve (b), where $r = 0.6$, hence the difference between curves (a) and (c) (the conventional curve) is larger than the difference between curves (b) and (c).

We now discuss the retrieval ability for pattern S′ with activity level r', when the threshold is set at the optimal threshold corresponding to S^1 with activity level r, i.e., $\theta = \theta_{\max}(r)$, [see (14)] but with $r \neq r'$. The equation that determines the final overlap $\langle m'(\infty)\rangle$ with pattern S′ can be obtained by a slight modification of (12):

$$\langle m'(t+1)\rangle = r'\{1 - 2\psi[\langle m'(t)\rangle - \theta_{\max}(r)]\} + (1-r')\{1 - 2\psi[\langle m'(t)\rangle + \theta_{\max}(r)]\}\,, \qquad (19)$$

In Fig. 2 we plot the fixed points of (19), which are the average final overlaps $\langle m'(\infty)\rangle$ between the state of the system and patterns S′ with activity level r', and when $\theta = \theta_{\max}(r)$, as a function of the noise level σ. We find that

$$\langle m'[t = \infty; \theta = \theta_{\max}(r)]\rangle < \langle m'[t = \infty; \theta = \theta_{\max}(r')]\rangle, \quad \text{when} \quad r' \neq r\,. \qquad (20)$$

The average final overlap with a pattern is, in general, reduced when the threshold of the neural network is set to the optimal threshold with respect to another pattern with a different activity level. For instance, in Fig. 2, curves (a) and (b) lie below curve (c). This difference is more pronounced when these two patterns (S′ and S^1) are biased differently, i.e., sign($r' - 0.5$) = − sign($r - 0.5$). In Fig. 2, the difference between curves (a) and (c) is greater than the difference between curves (b) and (c).

In summary of the results of this section, we find that the retrieval ability of the network with respect to a stored pattern, at a given noise level, can be improved significantly by setting the neuronal threshold to an *optimal threshold* given by (14), which is a function of the time, the activity level of the stored pattern, and the noise level, in comparison with the conventional zero-threshold case. After the threshold is set to $\theta_{\max}(t)$ with

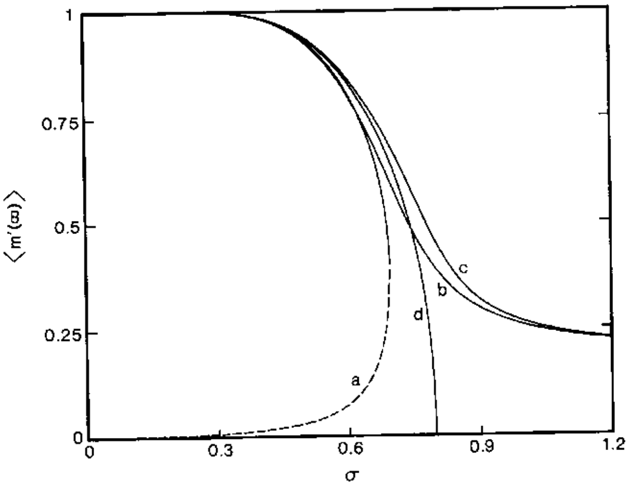

Fig. 2. Plot of average final overlap $\langle m'(t=\infty)\rangle$ between the state of the network and a stored pattern with activity level $r'=0.6$, vs. the standard deviation of the Gaussian noise σ, according to (19), while the network focuses its attention on stored patterns with activity level r by setting the neuronal threshold to $\theta=\theta_{\max}(r)$ given by (14): *(a)* $\theta=\theta_{\max}(r=0.4)$, *(b)* $\theta=\theta_{\max}(r=0.8)$, *(c)* $\theta=\theta_{\max}(r=r'=0.6)$. For comparison, we have also included curve *(d)*, in which the neuronal threshold $\theta=0$ and activity level r is arbitrary

respect to activity level r, however, retrieval abilities of the network with respect to stored patterns whose activity levels are different from r are, in general, not as good compared to patterns with activity level r. Hence choosing the optimal threshold $\theta_{\max}$ given in (14) may be a mechanism for the neural network to focus its *attention selectively* to a class of patterns with a particular activity level r. The effect of this kind of selective attention is more pronounced when the *attended* and *unattended* patterns are biased differently, i.e., one has a high activity level ($r>0.5$) and the other has a low activity level ($r'<0.5$) and vice versa.

There exist also other possible mechanisms of attention, e.g., through interactions between neural networks (Wang and Ross 1990a).

4 Effects of a constant threshold independent of pattern, time, and noise

In Sect. 3 we consider a particular dependence of the neuronal threshold on the time, the noise level, and the activity level and stored patterns, that optimizes the retrieval performance of the network with respect to a given pattern. We show in this section that if the threshold θ in (12) is chosen to be independent of variables such as the time t, the noise level σ, and the activity level r, the fixed points of the network display different behavior.

The following symmetry rules for y_1 defined in (16) (with $t=\infty$) are useful:

$$y_1[\langle m(\infty)\rangle, r, \theta, \sigma] = y_1[\langle m(\infty)\rangle, (1-r), -\theta, \sigma]\,, \quad (21)$$

$$y_1[-\langle m(\infty)\rangle, r, \theta, -\sigma] = y_1[\langle m(\infty)\rangle, (1-r), \theta, \sigma]\,, \quad (22)$$

and

$$y_1[-\langle m(\infty)\rangle, r, \theta, \sigma] = -y_1[\langle m(\infty)\rangle, r, -\theta, \sigma]\,. \quad (23)$$

We recall that $0\le r\equiv N_+/N\le 1$. Equation (21) means that for any $0\le r<0.5$, the fixed point of (12) remains the same if the activity level and the neuronal threshold are changed to $0.5\le r^*=1-r<1$ and $\theta^*=-\theta$, respectively. The fact enables us to focus only on cases where $0.5\le r<1$. As a check on this symmetry rule, we can easily verify by using (14) that $\theta_{\max}(r)=-\theta_{\max}(1-r)$.

Figure 3A is a plot of the average final overlap $\langle m(\infty)\rangle$ between the state of the neural network and pattern S^1 versus the standard deviation of the Gaussian noise, for a few typical cases where $\theta>0$ and $0.5\le r<1$ (see curves (a)–(c)), according to (12).

We see from Fig. 3A that hysteresis results from a constant threshold, even in the case of first-order interaction (for discussions on higher order interactions, see, e.g., Peretto and Niez 1986; Keeler et al., 1989; Wang and Ross 1990a, b). The dashed lines represent unstable fixed points, which define the domains of attraction of the stable fixed points (solid lines). Only the points above a given dashed line can be attracted to the corresponding solid line (stable fixed points). For a given neuronal threshold and at a given noise level, the initial overlap between the state of the network and pattern S^1 has to be greater than a critical value given by the point on the dashed line, in order for the network to converge to pattern S^1. When the initial condition of the neural network is below the dashed line, the network does not converge to pattern S^1 which is in the vicinity of the initial state of the system, instead it converges to the corresponding spurious state $-S^1$, according to (12). Hence the overlap between the final state of the system with pattern S^1 is negative in this case. The symmetry rule given by (23) shows that the absolute value of the average final overlap is equal to that with the sign of the threshold reversed ($-\theta<0$). Also the unstable (dashed) portions of curves (a)–(c) intersect with the $\langle m(\infty)\rangle$-axis at points $\langle m(\infty)\rangle=\theta$. Hence hysteresis exists for any positive $\theta<1$ and disappears only when $\theta=0$ (see curve (d)).

Compared to the zero-threshold case [curve (d)], the retrieval ability of the network, e.g., the average final overlap, the size of the basin of attraction, and the noise threshold $\sigma_c(\theta)$ above which there is no retrieval, is significantly reduced by the positive thresholds. In fact the more positive the threshold, the larger the reduction in the performance: the balloon-like curves ((c) through (a)) shrink to the upper left corner as θ increases towards 1. If $\theta\geqslant+1$, then on the average, the state of the network will not evolve towards the stored pattern that the network is initially close to, no specific stored pattern can be retrieved. We refer this condition as an anesthetized network. The network becomes completely incapable of retrieving when $\theta\ge+1$, i.e., $\sigma_c(\theta\ge+1)=0$, and the input signal fails to affect an anesthetized network. Hence, a positive threshold can desensitize ($0<\theta<1$) and anesthetize ($\theta\ge1$) the neural network with respect to patterns with high activities ($r>0.5$).

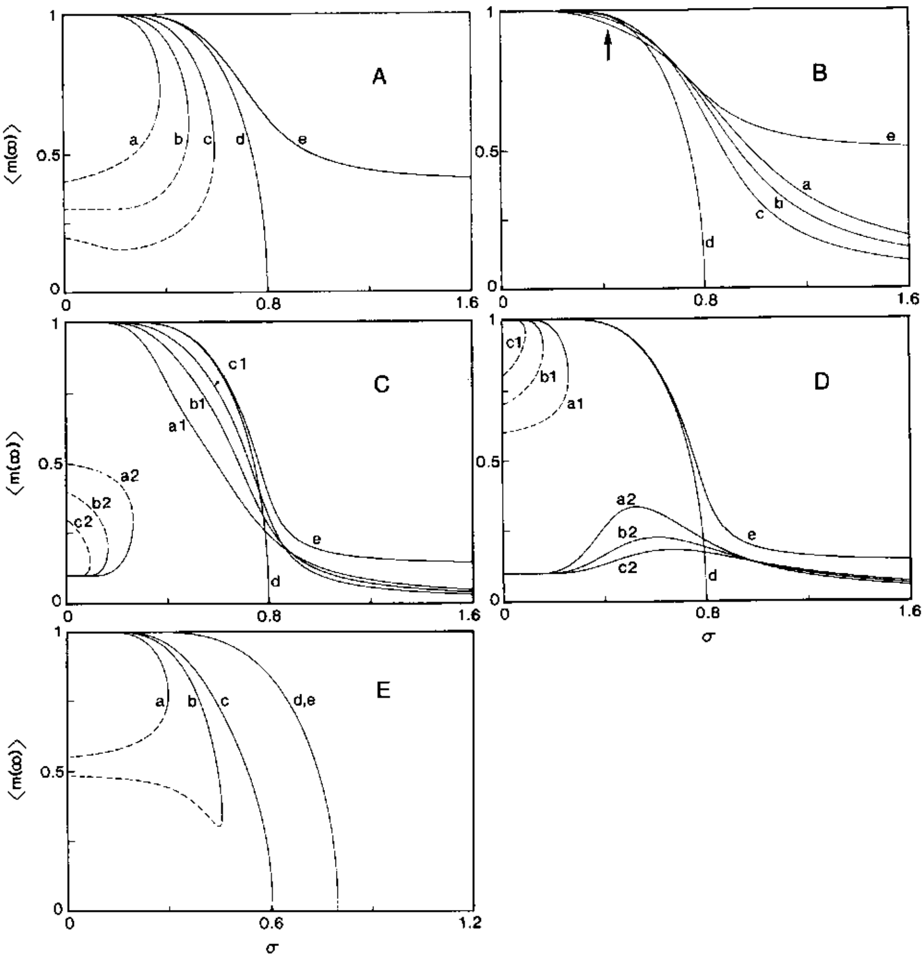

Fig. 3. A–E. Plot of average final overlap $\langle m(\infty)\rangle$ given by (12) vs. the standard deviation of the Gaussian noise σ for patterns with various activity levels. The thresholds here are independent of the noise level, the time, and the activity level. For comparison, we have also included *(d)*, in which the neuronal threshold $\theta = 0$ and activity level r is arbitrary, and curve *(e)*, in which $\theta = \theta_{\max}$. The dashed portions represent unstable fixed points which define the domains of attraction of the stable fixed points. **A** $r = 0.7$ *(a–e)* and $\theta > 0$ *(a–c)*. *(a)* $\theta = 0.4$ $(\theta > 1 - r = 0.3)$ *(b)* $\theta = 0.3$ $(\theta = 1 - r = 0.3)$ *(c)* $\theta = 0.2$ $(\theta < 1 - r = 0.3)$. At $\sigma = 0$, there are two solutions for $\langle m(\infty)\rangle$, i.e., 1 (stable) and θ (unstable). **B** $r = 0.75$ *(a–e)* and $-0.5 = (1 - 2r) \leq \theta < 0$ *(a–c)*. *(a)* $\theta = -0.4$ *(b)* $\theta = -0.3$ *(c)* $\theta = -0.2$. **C** $r = 0.55$ *(a–e)* and $-0.55 = -r \leq \theta < (1 - 2r) = -0.1$ *(a–c)*. *(a[illegible])* $\theta = -0.5$ *(b)* $\theta = -0.4$ *(c)* $\theta = -0.3$. At $\sigma = 0$, there are three solutions for $\langle m(\infty)\rangle$, i.e., 1 (stable), $|\theta|$ (unstable), and $|1 - 2r| = 0.1$ (stable). **D** $r = 0.55$ *(a–e)* and $\theta < -r = -0.55$ [*(a–c)*]. *(a)* $\theta = -0.6$ *(b)* $\theta = -0.7$ *(c)* $\theta = -0.8$. At $\sigma = 0$, there are three solutions for $\langle m(\infty)\rangle$, i.e., 1 (stable), $|\theta|$ (unstable), and $|1 - 2r| = 0.1$ (stable). **E** $r = 0.5$ *(a–e)*. *(a)* $\theta = 0.55$ *(b)* $\theta = -0.485$ *(c)* $\theta = 0.45$ *(d)* and *(e)* $\theta = \theta_{\max} = 0$

We want to show with some additional calculations some possibilities of improving retrieval, in the presence of noise, by varying the threshold of excitation. In cases where the thresholds are negative and $r > 0.5$, there exist three threshold regions, i.e., *(i)* $(1 - 2r) \leq \theta < 0$, *(ii)* $-r \leq \theta < (1 - 2r)$ and *(iii)* $\theta < -r$, where the network shows different behaviors, and the fixed points of (12) are shown in Fig. 3B, C and D, respectively. There is only one branch for the fixed point in case *(i)*, whereas bistable regions (hysteresis) appear in cases *(ii)* and *(iii)*. Again the dashed lines represent unstable fixed points, which define the domains of attraction of the stable fixed points (solid lines). We find that the retrieval ability of the network is improved by a positive threshold in the region of large noise level, in comparison with the zero-threshold case; however, the average final overlaps are reduced at low noise levels. In Fig. 3B, the final overlaps above the arrow are arranged in the following decreasing order: (e), (d), (c), (b), (a). These are the effects of a negative threshold (sensitization) to patterns with high activity levels.

For unbiased stored patterns ($r = 0.5$), (21) indicates that the average final overlap $\langle m(\infty)\rangle$ is unchanged if the sign of the threshold θ is reversed. We plot in Fig. 3E the average final overlap as a function of the noise level σ for the case with $r = 0.5$ and different choices of θ. Since the optimal threshold for unbiased patterns is $\theta_{\max} = 0$, according to (14), curve (e) ($\theta = \theta_{\max}$, and curve (d) ($\theta = 0$), the conventional curve, merge into one single curve. Figure 3E indicates that the larger $|\theta|$ is, the smaller $\langle m(\infty)\rangle$ and the noise threshold $\sigma_c(\theta)$. The network's retrieval ability with respect to unbiased stored patterns decreases as the magnitude of the neuronal threshold increases. In fact, $\langle m(\infty)\rangle = 0$ at any noise level σ, if $r = 0.5$ and $|\theta| \geq 1$. There is hysteresis in the $\langle m(\infty)\rangle$ vs. σ plots for $|\theta| \geq 0.484$.

Another interesting case occurs when $\theta = 0$, and $\langle m(\infty)\rangle$ becomes independent of r. If average final overlaps $\langle m(\infty)\rangle$ are plotted against the noise level σ for various activity levels at zero threshold, we obtain only one single curve, that is the conventional plot with $r = 0.5$ and $\theta = 0$ [curves (d) in Fig. 3].

5 Effects of a variable neuronal threshold on convergence rates

In the discussions above we have concentrated on the effects of a variable threshold on the fixed points of the neural network. The temporal dynamics of the system with a variable threshold can also be studied through

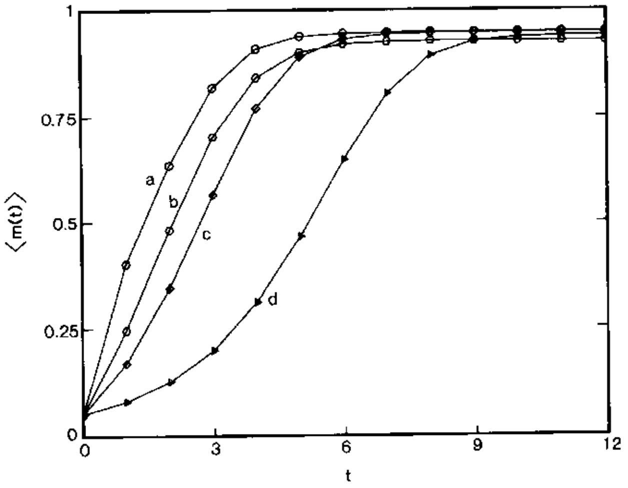

Fig. 4. Plot of average overlap $\langle m(t) \rangle$ given by (12) vs. time t (temporal dynamics) for various choices of the neuronal threshold: *(a)* $\theta = \theta_{\max}(r = 0.7)$, *(b)* $\theta = -0.3$, *(c)* $\theta = -0.15$, and *(d)* $\theta = 0$. The stored pattern to which the network converges has an activity level $r = 0.7$. The standard deviation of the Gaussian noise is $\sigma = 0.5$, and the initial condition is $\langle m(t = 0) \rangle = 0.05$. We compare the average final overlaps as follows: $\langle m(t = \infty;\ \theta = -0.3) \rangle < \langle m(t = \infty;\ \theta = 0) \rangle < \langle m(t = \infty;\ \theta = -0.15) \rangle < \langle m(t = \infty;\ \theta = \theta_{\max}) \rangle$

(12). For instance, in Fig. 4, we plot the average overlaps between the state of the network with pattern S^1 as functions of time for various choices of threshold. As already mentioned in Sect. 3, the convergence rate with $\theta = \theta_{\max}$ is the highest. Although the system converges faster with $\theta = -0.3$ than with $\theta = 0$, the average final overlap of $\theta = 0$ is greater.

6 Effects of a variable threshold on the memory capacity

In Sects. 5 and 6 we discussed two characteristics of the performance of a network, i.e., the average final overlap and the convergence rate, in the presences of a variable threshold. Now we study the effects of a variable threshold and an activity level distribution in the stored patterns on another important property of the network performance, the memory capacity.

If we let threshold $\theta = 0$, (12) reduces to (Kinzel 1985)

$$\langle m(t+1) \rangle = 1 - 2\psi[\langle m(t) \rangle]\,, \tag{24}$$

where $\psi(x)$ and σ are given by (11) and (7e), respectively. The fixed point $\langle m(\infty) \rangle$ of (24) can be calculated similarly. $\langle m(\infty) \rangle$ decreases as σ increases and vanishes for $\sigma \geq \sqrt{2/\pi}$. The memory capacity, i.e., the maximum number of pattern p that can be stored in the network, is qualitatively determined by letting $\sigma = \sqrt{2/\pi}$. By using the expression of given in (7e), we find that the storage capacity is greater when there exist either an activity level distribution or a deviation of the average activity level from 50%, or both (for similar conclusions in the absence of activity distribution, see Gardner 1988; Buhmann et al. 1989). This increase of memory capacity is due to the overlap between the stored patterns and does not imply an increase in the *information* capacity of the network. It is also evident from this analysis that the memory capacity increases with the size of the network N, since σ given in (7e) decreases as N increases.

If we consider a nonzero threshold, then for a given $\theta \neq 0$, the stability of a stored pattern depends on its activity level r. Hence a variable threshold has different effects on stored patterns with different activity levels. For instance, a positive threshold makes patterns with low activity levels ($r < 0.5$) stabler and makes patterns with high activity levels ($r > 0.5$) less stable. A pattern becomes unstable (disappears) when the average final overlap (see Figs. 2 and 3) vanishes. Hence as the threshold is varied, some stored patterns may disappear, which causes p to change. However, throughout this paper, we assume that the network is so large that the pattern of interest, S^1, never disappears and the effects of disappearances of other stored patterns are negligible when the threshold is varied.

7 Neurophysiological aspects related to a variable neuronal threshold

In this section we discuss the evidence of a variable neuronal threshold in neurophysiological systems and interesting phenomena caused by variations in thresholds. We begin by first relating the present model to neurophysiological systems. For references, see, e.g., Guyton 1981, chap. 2; Carpenter 1984, chap 2; and Kuffler et al. 1984, chap. 5.

When a real neuron is in its quiescent state, differences in the permeabilities and the initial concentrations of sodium and potassium ions (both extra- and intracellular) cause an electrical potential difference across the membrane, which is called the *resting potential* V_r and ranges normally from −95mV to −75mV, the negative sign meaning that the potential inside the membrane is lower than that outside the membrane. If this potential difference is made less negative (depolarized), e.g., by applying an external electric field, to a *firing potential* V_f, a voltage impulse (or action potential) is generated. Hence the threshold signal for a neuron to fire is

$$\theta = V_f - V_r\,. \tag{25}$$

We emphasize the difference between the definitions given in the present paper for the firing potential V_f and the threshold θ, since V_f is sometimes referred as "firing threshold" (Alberts et al. 1989, p. 1080).

The quantities used in the present model, i.e., the input $h_i(t)$, the state variables $S_i(t)$, the threshold θ_i, and the internal noise η_i, as well as its standard deviation σ_0, all correspond to electrical voltages in physiological neural systems or their artificial implementations. They are shifted and rescaled so that the ground state (V_r) and the excited (firing) state of a neuron are described by $S_i(t) = -1$ and $S_i(t) = +1$, respectively.

In such a representation, the maximum input signal, which is given by the first term of (6), is of the order of 1. The rest of the right hand side in (6) is noise, the internal interferences and external Gaussian noise, which are assumed to be not too large compared to the signal term. Hence qualitatively the state of *anesthesia* in the network corresponds to $\theta_i \geq 1$: the input signal into any neuron is always insufficient to overcome its threshold to achieve an action potential. The average of the "+1" and the "−1" states are 0: a positive θ_i makes it harder for the neuron to fire (desensitizition) and a negative θ_i make it easier for the neuron to fire (sensitizition). According to the scaling property given by (2b), overall synaptic changes can be represented by threshold changes. The maximum noise level a system with a zero-threshold can tolerate is $\sqrt{2/\pi} \approx 0.8$ (Keeler et al. 1989). The time t in the present model is scaled with the average length of an updating cycle. Equation (24) now simply means that the neuronal threshold can be varied by changing either the resting potential or the firing potential, or both, which can in turn be controlled by a number of variables, for example, permeabilities and concentrations of key ions such as K^+, Na^+, Ca^{++} (Hodgkin and Huxley 1952), temperature, and other chemicals, to be discussed.

In general any condition that favors the potassium (sodium) transport across the membrane, rather than the sodium (potassium) transport, tends to raise (reduce) the neuronal threshold (Carpenter 1984, p. 40). For instance, the drug veratrine increases the membrane permeability to sodium, and the lowest strength of stimulus needed to elicit an impulse is significantly reduced by applying this drug. Sometimes it can make the fiber so excitable that it fires spontaneously without any extraneous excitation. A lower-than-average calcium ion concentration in the extracellular fluids has the same effects on a neuron. For patients who have lost their parathyroid glands and who therefore cannot maintain normal calcium ion concentrations, this condition may cause spontaneous respiratory muscle spasm and can be fatal. On the other hand, a high calcium concentration increases neuronal threshold and makes a neuron less excitable. There exist many other such "stabilizers", some of which are used as local *anesthetics*, e.g., procaine, cocaine, tetracaine. When the concentrations of these drugs are large enough, the neuronal threshold can be raised so high that the maximum signal received by a neuron is always below its firing threshold and the neuron thus never responds to signals of action potentials from other neurons, which stops nerve impulses from affecting the *anesthetized* region (Guyton 1981, p. 28). Changes in neural excitability can also be caused by damage to sensory fibers, which produces reduced (increased) sensation and is also called hype(re)sthesia (Willis and Grossman 1977, p. 410). An overall change (either increase or decrease) in neuronal threshold, e.g., caused by special chemicals and physical injuries discussed above, is related to the *constant threshold* discussed in the present paper.

It has been known that cognitive factors, such as the attentional state of a higher animal or the significance of an event, can alter the excitability of sensory neurons. For instance, athletes and soldiers frequently fail to notice painful injuries until after the game or battle. Human subjects asked to focus their attention elsewhere while receiving a painful stimulus rate the pain as less than that experienced when they are allowed to attend to the pain (Leventhal et al. 1979). The following experiment (Bushnell et al. 1984, 1987) is particularly interesting. Single unit activity was recorded in thermally sensitive neurons of awake monkeys whose attentional states were controlled by varying the tasks that they performed for a juice reward. These tasks involve detecting changes in visual or thermal stimuli. In some instances, the monkeys received a reward for detecting changes in the temperature of a thermode positioned on the face, and thus were encouraged to attend to the thermal sensation. At other times, however, they received their award for detecting changes in a visual cue, which distracted attention from changes in temperature of the thermode. It was found that the thermally sensitive neurons were much more excitable when a monkey was encouraged to attend to the thermal stimulus in comparison with situations in which the monkey was encouraged to attend elsewhere (visual cue in this experiment). This "*task-related*" modulation in neuronal excitability (also see Desmedt 1977, 1979) is closely analogous to our *pattern-dependent* neuronal threshold change proposed in the present paper (see 14), whereas the *constant* threshold shift is similar to the effects of chemicals discussed in previous paragraphs.

(De)sensitization and anesthesia are concepts usually confined to sensory neurons. In the present paper, we have incorporated the concepts of (de)sensitization and anesthesia into associative neural systems. By considering an associative neural network model we have shown that the system can achieve such higher functions as selective attention by purposefully changing its neuronal threshold in a prescribed manner; (de)sensitization in an associative neural network causes a variety of interesting effects on stationary points and dynamical properties (such as relaxation time) of the system.

The neuronal threshold can also change according to the dynamics of the neuron. For example, the threshold is considerably *increased* when the neuron is repetively stimulated *below* its firing potential V_f. A slow depolarization of the membrane has the same effect. This phenomenon is called *accommodation* (Guyton 1981, p. 27). In fact, if the rate of depolarization is too slow, the neuron may never fire at all, no matter how much it is depolarized, i.e., the threshold tends to infinity (Fabre 1927). Hill (1936) has developed a model for accommodation, in which a variable threshold is introduced. The Hill model has been tested against detailed experiments and good agreement has been found (e.g., Bradley and Somjen 1961).

Opposite to accommodation is *potentiation* (Carpenter 1984, p. 76): the neuronal threshold *decreases* when the neuron is repetetively stimulated *above* the firing potential V_f. This type of threshold change is in the spirit of hysteresis in a single neuron proposed recently by Hoffmann (Hoffmann 1986; also see Hoffmann et al. 1986) and studied in detail in associative neural network models by Wang and Ross (1990a). Neuronal hysteresis represents an increase in excitability (or decrease in threshold) for firing neurons and no change for quiescent neurons. Hence neuronal hysteresis introduces a threshold which is *neuron-specific* dependent on the firing history. In contrast, our present discussions concentrate on various cases where thresholds are identical for all neurons, although (10) is more general in that the neuronal thresholds there are arbitrary.

Acknowledgements. We thank the referee for a most careful reading of the paper and for many helpful questions and comments which improved the paper significantly. We thank Bill Vance for his help with numerical algorithms. This work was supported in part by the National Science Foundation.

Appendix: proof of (10)

We adapt a mean-field approximation that the average can be taken inside the sign function, i.e., $\langle \mathrm{sign}(x)\rangle \approx \mathrm{sign}(\langle x\rangle)$, and rewrite (8) as follows

$$\langle m(t+1)\rangle = \frac{1}{N}\sum_{i=1}^{N} A_i \,, \tag{A1}$$

with

$$A_i = S_i\,\mathrm{sign}[S_i\langle m(t)\rangle - \theta_i + \eta'] \,. \tag{A2}$$

We consider *(i)* $S_i = +1$ and *(ii)* $S_i = -1$ separately. When $S_i = +1$, (A2) gives

$$A_i = P_+ - P_- \,, \tag{A3}$$

where $P_\pm$ are the probabilities that $\mathrm{sign}[\langle m(t)\rangle - \theta_i + \eta'] = \pm 1$, respectively. Since η' is a Gaussian noise, we have $\mathrm{sign}[\langle m(t)\rangle - \theta_i + \eta'] = -1$, if $\eta' < -[\langle m(t)\rangle - \theta_i]$ Considering the fact that $P_+ + P_- = 1$, we obtain

$$\begin{aligned} A_i &= 1 - 2P_- \\ &= 1 - 2\frac{1}{\sqrt{2\pi}}\int_{-[\langle m(t)\rangle - \theta_i)/\sigma}^{-\infty} \mathrm{e}^{-x^2/2}\,\mathrm{d}x \\ &= 1 - 2\psi[\langle m(t)\rangle - \theta_i]\}, \quad \text{if} \quad S_i = +1 \,, \end{aligned} \tag{A4}$$

where ψ is defined by (11).

In the $S_i = -1$ case, we have, instead of (A3),

$$A_i = -P_+ + P_- = 1 - 2P_+ \,, \tag{A5}$$

where $P_\pm$ are the probabilities that $\mathrm{sign}[-\langle m(t)\rangle - \theta_i + \eta'] = \pm 1$, respectively. Hence

$$A_i = 1 - 2\psi\{[\langle m(t)\rangle + \theta_i]\}, \quad \text{if} \quad S_i = -1 \,. \tag{A6}$$

Combination of (A1), (A4) and (A6) gives (10).

References

Abelles M (1982 Local cortical circuits. Springer, New York Berlin Heidelberg, p21

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1989) Molecular biology of the cell, 2nd edn. Garland, New York

Amit DJ, Gutfreund H, Sompolinsky H (1985) Spin-glass models of neural networks. Phys Rev A32:1007–1018

Amit DJ, Gutfreund H, Sompolinsky H (1987a) Information storage in neural networks with low levels of activity. Phys Rev A35:2293–2303

Amit DJ, Gutfreund H, Sompolinsky H (1987b) Statistical mechanics of neural networks near satuation. Ann Phys (NY) 173:30–67

Bradley K, Somjen GG (1961) Accommodation in motoneurons of the rat and the cat. J Physiol 156:75–92

Buhmann J, Divko R, Schulten K (1989) Associate memory with high information content. Phys Rev A39:2689–2692

Bushnell MC, Duncan GH, Dubner R, He LF (1984) Activity of trigeminothalamic neurons in medullary dorsal horn of awake monkeys trained in a thermal discrimination task. J Neurophysiol 52:170–187

Bushnell MC, Duncan GH, Lund JP (1987) Gating of sensory transmission in the trigeminal system. In: Wise SP (ed) Higher brain functions. Wiley, New York

Carpenter RHS (1984) Neurophysiology. University Park Press, Baltimore, Md

Cortes C, Krogh A, Hertz JA (1987) Hierarchical associative networks. J Phys A20:4449–4455

Crick F (1989) The recent excitement about neural networks. Nature 337:129–132

Derrida B, Gardner E, Zippelius A (1987) An exactly solvable asymmetric neural network model. Europhys Lett 4: 167–173

Desmedt JE (ed) (1977) Attention, voluntary contraction and event-related cerebral potentials. Karger, New York; (1979) Cognitive components in cerebral event-related potentials and selective attention. Karger, New York

Domany E (1988) Neural networks: a biased overview. J Statist Phys 51:743–775

Fabre P (1927) L'excitation neuro-musculaire par les courants progressifs chez l'homme. C R Acad Sci (Paris) 184:699–701

Gardner E (1988) The space of interactions in neural network models. J Phys A21: 257–270

Grossberg S (1970) Neural pattern discrimination. J Theor Biol 27:291–337

Gutfreund H (1988) Neural networks with hierarchically correlated patterns. Phys Rev A37:570–577

Guyton AC (1981) Basic human neurophysiology. Saunders, Philadelphia

Hebb DO (1949) The organization of behavior. Wiley, New York, p 44

Hill AV (1936) Excitation and accommodation in nerve. Proc R Soc (London) Ser B 119:305–355

Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol 117:500–544

Hoffmann GW (1986) A neural network model based on the analogy with the immune system. J Theor Biol 122:33–67

Hoffmann GW, Benson MW, Bree GM, Kinahan P (1986) A teachable neural network model based on an unorthodox neuron. Physica 22D:233–246

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

Hopfield JJ (1984) Neurons with graded response have collective computational properties like those of two-state neurons. Proc Natl Acad Sci USA 81:3088–3092

Keeler JD, Pichler EE, Ross J (1989) Noise in neural networks: Thresholds, hysteresis, and neuromodulation of signal-to-noise. Proc Natl Acad Sci USA 86:1712–1716

Kinzel W (1985) Learning and pattern recognition in spin glass models. Z Phys B-Condensed matter 60:205–213

Kohonen T (1988) Self organization and associative memory. Springer Series in Information Science, vol 8, 2nd edn. Springer, Berlin Heidelberg New York

Krogh A, Hertz JA (1988) Mean-field analysis of hierarchical associative networks with 'magnetisation'. J Phys A21: 2211–2224

Kuffler S, Nicholls JG, Martin AR (1984) From neuron to brain, 2nd edn. Sinauer, Sunderland, Mass

Leventhal H, Brown D, Shacham S, Engquist G (1979) Effects of preparatory information about sensations, threat of pain, and attention on cold pressor distress. J Personal Social Psychol 37:688–714

Little WA (1974) The existence of persistent states in the brain. Math Biosci 19:101–120

McCulloch WS, Pitts W, Bull (1943) A logical calculus of the ideas immanent in nervous activity. Math Biophys 5:115–133

Noest AJ (1988) Phasor neural networks. In: Anderson DZ (ed) Neural information processing system. American Institute of Physics, New York, pp 584–591

Parga N, Virasoro MA (1986) The ultrametric organization of memories in a neural network. J Phys (Paris) 47:1857–1864

Peretto P, Niez JJ (1986) Stochastic dynamics of neural works. IEEE Trans SMC-16:73–83

Sompolinsky H (1988) Statistical mechanics of neural works. Phys Today 41:70–80

Wang L, Ross J (1990a) Synchronous neural networks of non-linear threshold elements with hysteresis. Proc Natl Acad Sci (USA) 87:988–992

Wang L, Ross J (1990b) Interactions of neural networks: models for distraction and concentration. Proc Natl Acad Sci (USA) 87:7110–7114.

Willis WD Jr, Grossman RG (1977) Medical neurobiology, 2nd edn. Mosby, St. Louis

Prof. John Ross
Department of Chemistry, Stanford University
Stanford, CA 94305, USA