*Proc. Natl. Acad. Sci. USA*
Vol. 87, pp. 9241–9245, December 1990
Neurobiology

# Modulation of function and gated learning in a network memory

**(model neuron/associative memory/neuromodulation/synaptic plasticity)**

L. F. ABBOTT

Physics Department and Center for Complex Systems, Brandeis University, Waltham, MA 02254

*Communicated by John Ross, July 17, 1990 (received for review April 4, 1990)*

**ABSTRACT** **Memory and learning are studied in a model neural network made from component cells with a variety of realistic intrinsic dynamic behaviors. Modulation of intrinsic cellular characteristics causes a network to switch between two entirely different modes of operation. In one mode the network acts as a selective, long-term associative memory, whereas in the other it is a nonselective, short-term latching memory. Such functional modulation can be used as a mechanism for initiating and terminating learning in a network associative memory.**

In addition to displaying complex collective phenomena in networks, biological neurons exhibit a wide variety of intrinsic dynamical behaviors, even in isolation. How does intrinsic neuronal dynamics affect network behavior and how can modulation of intrinsic cellular properties be used to modify and control network function? To address these questions, a model neuron has been constructed, which displays realistic dynamic behavior and yet is simple enough to be amenable to detailed mathematical analysis and computer simulation (1). I previously demonstrated (1) how a network of these dynamic model cells can act as an associative memory with storage and recall properties similar to those of well-known attractor network models (2–4) (for reviews, see refs. 5 and 6). Here, I show how modulation of intrinsic cellular properties can be used to alter and control network function and to initiate and terminate learning in a network associative memory.

Attractor network models have provided a fertile testing ground for ideas about memory and learning. However, the binary model neuron, which forms the basis of these networks, is essentially devoid of intrinsic dynamics. To study the impact of intrinsic behavior and the role of neuromodulation, a richer model of the basic cell must be constructed. The model neuron used here is based on a solution of a piecewise linear form of the FitzHugh–Nagumo equations (7, 8) but is formulated much like attractor models (2–4), especially those with time-dependent thresholds (9) and hysteresis (10). The model cells display realistic intrinsic behaviors, such as postinhibitory rebound, postburst hyperpolarization, fatigue, intrinsic oscillation, plateau behavior, and bistability (1).

The behavior exhibited by a given model neuron depends of course on the values of various model parameters. A particularly interesting set of parameter values leads to the steady-state characteristics shown in Fig. 1. Four different behaviors are possible depending on the total synaptic current entering the cell, $I$, and on the value of an additional model parameter, $a$. Not surprisingly the cell hyperpolarizes under the influence of negative current and depolarizes for positive current. In addition two more complex phenomena are observed. In the absence of synaptic current, the cell oscillates when $0.5 < a < 1$, whereas for $0 < a < 0.5$ it is bistable. Fig. 1 *B* and *C* shows both oscillation and bistability. Since individual action potentials are not modeled, this figure shows a smoothed membrane potential with spikes removed. However, when the model cell is depolarized, the average membrane potential can serve as a measure of the spiking rate. The waveform for an oscillating model cell is shown in Fig. 1*B*. In the bistable mode, the model cell can remain indefinitely in either a depolarized or a hyperpolarized state. A brief positive synaptic current pulse will place the cell in the depolarized state, whereas a negative pulse will cause it to remain hyperpolarized as shown in Fig. 1*C*.

Plateau behavior, bistability, and oscillation are closely related phenomena. It is likely that many cells exhibiting any one of these can be induced through modulation to display the others as well. Typically, a cell must have a negative resistance region in its fast current-vs.-voltage ($I$–$V$) curve to oscillate. If the negative resistance characteristics are sufficiently enhanced by some modulatory substance, the cell will become bistable.

Many neuromodulators capable of changing cellular $I$–$V$ characteristics have been identified in biological preparations (11, 12). These can be applied either globally to an entire network or locally to a particular neuron at a synaptic junction. Here I will study global neuromodulation, concentrating on the effects of a hypothetical neuromodulator that enhances the negative resistance characteristics of network cells. These effects are simulated in the model by changing the value of the parameter $a$. In the following sections I will show (*i*) how a network of oscillatory cells can act as a selective, long-term, associative memory; (*ii*) how a modulator that makes these cells bistable can modify the function of this network, making it a nonselective, short-term, latching memory; and (*iii*) how this switching can be used to gate the learning process in a network associative memory.

## A Network of Dynamic Model Neurons

The dynamic behavior of any neuronal cell depends on the voltage and time dependences of its membrane conductance channels. To deal with large networks of model cells, we must provide a compact and manageable approximate description of membrane currents, which is nevertheless accurate enough to produce interesting intrinsic dynamics. To do this, we will not model individual action potential spikes but will consider instead a spike-averaged membrane potential. When the cell is firing, this potential can be used as a measure of the spiking rate. The dynamic model cell used here is based on a well-known approach (7, 8). The total membrane current is divided into two pieces, one consisting of all currents that respond rapidly, say in a time on the order of a few msec or less, and another composed of the slow currents responding in times on the order of tens of msec. The fast set of currents is then considered to be an instantaneous function of the spike-averaged membrane potential. Let $f(v)$ be the fast outward membrane current at cell potential $v$ and let $u$ be the slow component of the membrane current. All currents and potentials are in arbitrary units and can be rescaled for

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: $I$–$V$, current vs. voltage.

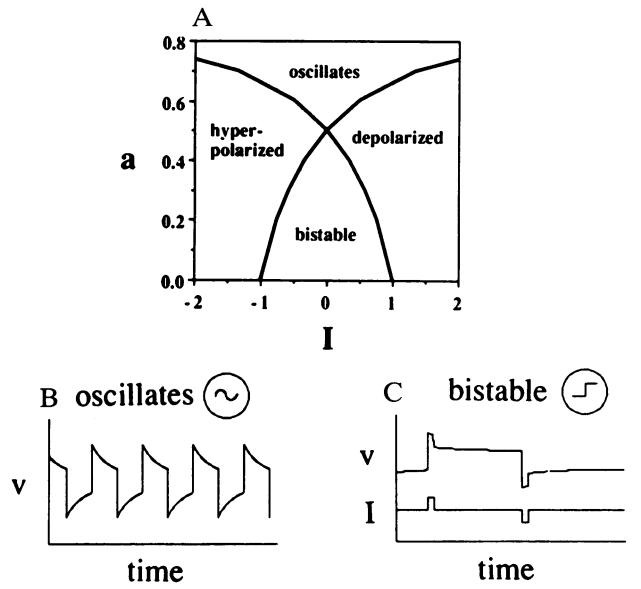


FIG. 1. Phase diagram for a model neuron. (*A*) $I$ is the total synaptic current entering the cell and $a$ is a model parameter. With $I = 0$, the cell oscillates when $a$ is >0.5 and is bistable for $a$ <0.5. The spike-averaged cell potential as a function of time in the oscillatory (*B*) and bistable modes (*C*) is shown. In the bistable mode (*C*), the cell potential (upper trace) can remain either depolarized or hyperpolarized indefinitely, and brief current pulses (lower trace) can cause transitions between these two stable states.

specific applications. It is assumed that the cell capacitance is small enough so that on the 10-msec time scale of interest the total current leaving the cell must match that coming in,

$$f(v) + u = I, \qquad [1]$$

where $I$ is the total synaptic and external current entering the cell. The current component $u$ responds slowly to changes in the membrane potential, and its behavior has traditionally been modeled (7, 8) by using a linear first-order differential equation

$$\tau \frac{du}{dt} = av - (1 - a)u. \qquad [2]$$

The parameter $a$ in this equation is the same as the $a$ of Fig. 1. The parameter $\tau$ sets the time scale for intrinsic dynamical behavior such as oscillation. For my purposes this time scale is taken to be on the order of ten or tens of msec.

For the purposes of this paper, it will suffice to use a piecewise linear form for the fast current $f(v)$, consisting of two positive resistance regions for hyperpolarized potentials less than $v = -1$ and depolarized potentials greater than $v = +1$ connected by an intermediate negative resistance region. The choice of the points $v = \pm 1$ is of course arbitrary, and these may be moved to any desired values by shifting and rescaling $v$. The exact form of the negative resistance region is not important, but to keep track of the two positive resistance regions we define a binary variable $S$ satisfying

$$S = \begin{cases} +1 \text{ if } v \geq 1 \\ -1 \text{ if } v \leq -1 \end{cases}. \qquad [3]$$

We do not need to define $S$ in the region $|v| < 1$ because in the limit of small cell capacitance the membrane potential jumps instantaneously over this region, always remaining $> 1$ in absolute value. In the relevant regions $v \leq -1$ and $v \geq 1$, the rapid component of the membrane current can then be written as

$$f(v) = v - 2S, \qquad [4]$$

and Eq. 1 can be solved for the spike-averaged potential

$$v = I + 2S - u. \qquad [5]$$

By using this result, Eq. 3 can be written more compactly as

$$S = \text{sgn}[S + I - u]. \qquad [6]$$

A network of $N$ model neurons is characterized by $N$ binary variables $S_i = \pm 1$ *with* $i = 1 \ldots N$, which keep track of which branch of the fast *I–V* curve is being used for cell $i$. Thus, if cell $i$ is hyperpolarized and silent, $S_i = -1$, whereas if it is depolarized and firing, $S_i = +1$. Time is divided up into units of $\approx 1$ msec, and the state of cell $i$ at time $t + 1$ is given in terms of its state at time $t$ by the updating rule

$$S_i(t + 1) = \text{sgn}[S_i(t) + I_i(t) - u_i(t)], \qquad [7]$$

which assures that Eq. 6 is satisfied for each cell of the network. $I_i(t)$ is the total synaptic current entering cell $i$ at time $t$. As in attractor models, a matrix element $J_{ij}$ characterizes the synaptic coupling between neuron $j$ and neuron $i$. The sign of $J_{ij}$ governs whether this synapse is excitatory or inhibitory, and its magnitude determines the strength of the coupling. The total synaptic current entering cell $i$ is (up to a constant, which can be absorbed into the definition of $u$)

$$I_i(t) = \sum_{j=1}^{N} J_{ij} S_j(t). \qquad [8]$$

The function $u_i(t)$ representing the slow component of the cell membrane current is determined by an updating rule obtained by integrating Eq. 2 over one time step

$$u_i(t + 1) = u_i(t)e^{-1/\tau_i} + a(I_i(t) + 2S_i(t))(1 - e^{-1/\tau_i}). \qquad [9]$$

For simplicity, in this discussion all cells of the network have been given the same value of $a$, but they are allowed different intrinsic time constants $\tau_i$.

Fig. 2 shows three relevant behaviors exhibited by a network of model cells. For clarity a network of only five cells is shown (a much larger network is discussed in the next section). In Fig. 2 *A* and *B*, the model cells are intrinsically oscillatory, whereas in Fig. 2*C* they are bistable. In all three cases the network is initialized by a set of brief current pulses entering each cell. These pulses place the network in an initial state characterized by a specific pattern of firing and silent cells. In Fig. 2*A* the model cells are uncoupled, so after being initialized they oscillate incoherently, each at its own particular intrinsic frequency.

Fig. 2*B* shows coherent oscillation, which forms the basis of associative memory in the model (1). In an associative memory, the synaptic couplings are adjusted (through learning) to pick out certain patterns of network activity as memory patterns. If the pattern of activity initiated by the input is close to one of the memory patterns, the network will respond by phase locking (13–15). As in Fig. 2*B*, the response of the network to a recognized input will be coherent oscillation between the memory pattern recalled by that particular input and its inverse (obtained by making all passive cells active and all active cells passive). Coherent oscillation allows for the recovery of stored memory patterns. Phase locking depends both on the specific form of the synaptic couplings and on the nature of the input state. Because of this, the memory is selective; only inputs close to one of the stored memory patterns will be recognized. In addition, the

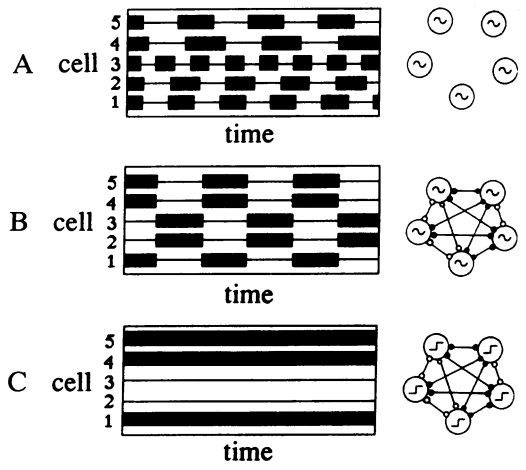


FIG. 2. Three modes of operation for a five-cell network. The horizontal axis is time and the vertical axis gives the number of the cell being plotted. The depolarized or firing state and the hyperpolarized or silent state are denoted by the presence or absence of a black band, respectively. To begin the simulations an initial pattern of activity is imposed. In *A* the cells are uncoupled so they oscillate independently at their own intrinsic frequencies. Hebb-type couplings have been used in *B* to produce phase locking. The cells oscillate in unison at a common frequency between the memory pattern and its inverse. For *C* the variable *a* has been reduced so that the cells are bistable. The network maintains whatever pattern of activity was imposed initially. Circuit diagrams show the different network configurations; open circles represent excitatory synapses, and solid circles represent inhibitory synapses.

memory is long-term since memory patterns will be retained as long as synaptic couplings are not modified.

Whether or not a given input state will converge to a coherent oscillating pattern depends on the basin of attraction of the limit-cycle attractor associated with the appropriate memory pattern. The basin of attraction defines how similar the input state has to be to a memory pattern in order to be recognized and mapped to it. Basins of attraction in the oscillatory model appear to be quite similar to those in fixed-point attractor networks, but it should be noted that in both the oscillating and fixed-point cases much remains to be learned about the nature of the attraction basins.

Phase locking is frequently seen in coupled oscillatory systems. However, the particular nonlinear oscillator being discussed here exhibits an extremely robust form of synchronous behavior in comparison with, for example, phase-coupled oscillators (13–15). This is due to the fact that the present model allows a driven oscillator to alter both the size of its phase-space limit cycle and the rate of motion along the limit cycle to facilitate synchronization with the driving current.

In Fig. 2*C* the parameter *a* has been decreased, causing the network cells to become bistable. This simulates the effect of a globally applied neuromodulator. As a result, the network maintains a fixed pattern of activity identical to that imposed initially by the input current pulses. In this mode *any* input presented to the network will result in a fixed pattern of network activity identical to the initial state. The network therefore functions as a nonselective, short-term, latching memory. If the value of *a* is increased back to its original value, simulating the clearing of the neuromodulator, the type of behavior shown in Fig. 2*B* will resume.

Classes of behavior other than those shown in Fig. 2 can also be seen in network simulations. For example phase locking in various ratios other than 1:1 and chaotic behavior have been observed. These are discussed more fully in ref. 1 but will not be discussed here as they play no role in the memory and learning tasks being considered.

### Gated Learning

The behavior shown in Fig. 2 *B* and *C* can be used to produce either long-term, associative or short-term, latching memory. A particularly interesting application of the ability of networks of dynamic model cells to switch between these two functions is gated learning. Learning is the process by which network connections appropriate for performing a given task are established. In oscillatory networks, as in attractor networks, learning is accomplished through modification of synaptic strengths (long-term potentiation) (16–26) in response to the activity of pre- and/or postsynaptic cells. The learning process requires that a fixed pattern of network activity be maintained long enough for significant synaptic modification to take place. However, while a network is acting as an associative memory, uncontrolled synaptic changes cannot be allowed since this could result in loss of stored information. Therefore learning must be gated. In conventional attractor networks this is done by assuming that external latching circuitry is available and by regulating synaptic plasticity. In the example considered here, the oscillating associative memory can act as its own latch. Furthermore, it will not be necessary to regulate the plasticity of the synapses. Learning can be initiated and terminated solely by modulating intrinsic cellular characteristics.

This form of gated learning relies on a key assumption about a temporal threshold for synaptic plasticity. We assume that synaptic modification requires that a fixed pattern of pre- and postsynaptic activity be maintained over a relatively long (on the order of 100 msec) period of time. For example, in order for Hebb-type modifications to take place, both the pre- and postsynaptic cells must fire continuously for on the order of 100 msec. Such a temporal threshold assures that no unwanted synaptic modification will take place while the oscillatory network is acting as an associative memory because in this case cells continually oscillate between active and passive states with periods on the order of tens of msec (see Fig. 2*B*). However, when learning *is* desired, synaptic modification can be induced by modulating the network so that it becomes a short-term, latching memory. Then any subsequent input will lead to a fixed pattern of network activity (see Fig. 2*C*) and, after the required 100 msec or so, synaptic modification will occur.

Memory recovery and gated learning in a 100-cell network are illustrated in Fig. 3. In preparation for this figure, a 100 × 100 synaptic coupling matrix encoding nine different patterns of activity was constructed using a generalized Hebb learning rule. Specifically, denoting the nine patterns by $\xi_i^\mu = \pm 1$ with $i = 1 \ldots N$ and $\mu = 1 \ldots 9$, the coupling matrix was taken to be

$$J_{ij} = \frac{1}{4N} \sum_{\mu=1}^{9} \xi_i^\mu \xi_j^\mu. \qquad [10]$$

One of the learned patterns was taken to be all cells firing at once ($\xi_i = 1$ for all $i$). Recall of this pattern by the network is shown in Fig. 3*A*. When the all-firing pattern is imposed as an initial state, the network phase locks, oscillating coherently between this pattern and its inverse. At this point, some readout mechanism similar to what is needed in ordinary attractor neural networks could be used to extract the recognized pattern. This mechanism is not included as a part of the present model. Another pattern consisting of alternating regions of firing and silent cells is used as the input state in Fig. 3*B*. The network has not yet learned this pattern, so it is not recognized, and the result is a pattern of activity completely unrelated to the initial input.

Fig. 3*C* shows learning. The same input pattern used in Fig. 3*B* is initially imposed on the network, but now the cells are

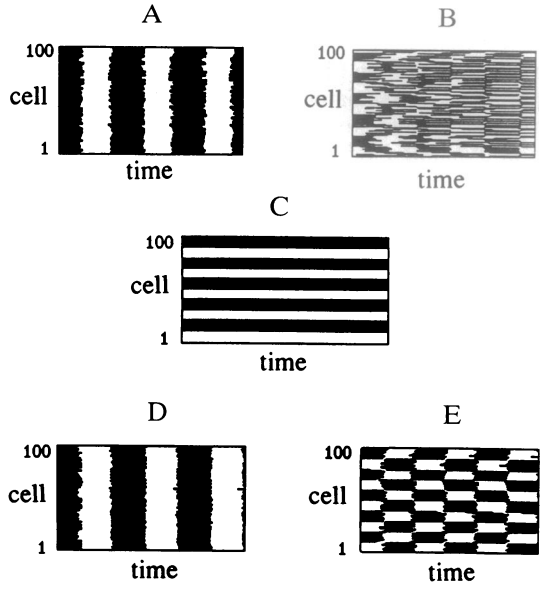


FIG. 3. Learning in a model network. A network of 100 model cells is shown using the same conventions and procedures as in Fig. 2. To compress the figure, the spaces between cell traces and individual cell baselines have been removed. (*A*) Recovery of a learned pattern (all cells firing) through phase locking. (*B*) An initial pattern of activity, which has not been learned yet, results in behavior unrelated to the input. (*C*) Modulation of cellular characteristics results in latching behavior, and the same input used in *B* now leads to a fixed pattern of network activity. This induces synaptic modification governed by the Hebb learning rule. (*D*) When the network cells regain their oscillatory characteristics, the pattern recognized in *A* is still recognized. (*E*) In addition, the network now recognizes the pattern it did not identify in *B* due to the learning that occurred in *C*.

bistable rather than oscillatory because the parameter $a$ has been reduced from $a = 0.6$ to $a = 0.1$. As a result, the network maintains this particular pattern of activity, and synaptic modification, which did not occur in Fig. 3 *A* and *B* due to the continual oscillations, now takes place. The only modification made to initiate and terminate learning in this example is a change in the value of the parameter $a$ simulating the effect of a modulatory substance. Once the threshold condition is satisfied, learning is accomplished with the same rule used originally to store the initial nine patterns. That is, after the threshold period, while the network is latched in the state $S_i$, the synaptic couplings are modified by the rule

$$J_{ij} \rightarrow J_{ij} + \frac{1}{4N} S_i S_j. \qquad [11]$$

Although this rule was used for the results being shown, any of a host of other synaptic modification rules (6) [including more realistic ones (27)] could be employed just as well.

When $a$ is increased back to 0.6 (simulating the clearing of the modulator) and the cells become oscillatory again, the original pattern used in Fig. 3*A* is still recognized, as seen in Fig. 3*D*. In addition, because of the synaptic modification that occurred during learning, Fig. 3*C*, the pattern that was not identified in Fig. 3*B* is now recognized and results in a coherent pattern of oscillatory activity, Fig. 3*E*. (The two patterns used were chosen for visual clarity. Any uncorrelated patterns could have been used. The 100 oscillators were given random intrinsic periods that varied by 50% over the population, and the average value of the $\tau_i$ was 25.) It should be stressed that the only modification made to initiate and terminate learning was a change in the value of the parameter $a$. In particular the synaptic modification rule (Eq. **11**) is applicable at all times to be performed whenever the temporal threshold condition is satisfied.

## Discussion

Fig. 3 shows clearly that a network of oscillators can exhibit memory recall through phase locking as well as modulation of function and gated learning through modification of intrinsic cellular characteristics. There is both speculation and experimental evidence that oscillating or bursting neurons and their temporal correlations play a significant role in various perceptual processes (28–39). The use of oscillatory rather than tonically firing elements has obvious advantages. The temporal information contained in an oscillating pattern of activity can be used for timing information, and phase locking of oscillating subgroups can be used to label related patterns. This idea has been discussed in other oscillatory models (ref. 41; D. Horn and M. Usher, personal communication) and is also being incorporated into the model presented here.

In the discussion up to now I have taken the position that the oscillators from which the network is constructed represent individual neurons. However, this is not the only possible interpretation. It has been shown (40) that populations of excitatory and inhibitory neurons can exhibit oscillation and bistability as collective phenomena. This group behavior can be treated mathematically much as the analogous neuronal behavior was treated here. Therefore a possible alternate interpretation of the present work is that the individual oscillators of the network are actually themselves populations of nonoscillatory neurons. Of course in some cases oscillations in neuronal populations may result from both individual and collective phenomena.

I am extremely grateful to Eve Marder, Jorge Golowasch, Frank Buchholtz, and John Lisman for help, advice, and encouragement. This research was supported in part by Department of Energy Contract DE-AC0276-ER03230.

1. Abbott, L. F. (1990) *J. Phys. A*, **23**, 3835–3859.
2. McCulloch, W. S. & Pitts, W. (1943) *Bull. Math. Biophys.* **5**, 115–133.
3. Little, W. A. (1975) *Math Biosci.* **19**, 101–120.
4. Hopfield, J. J. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558.
5. Amit, D. (1989) *Modelling Brain Function: The World of Attractor Neural Networks* (Cambridge Univ. Press, Cambridge, U.K.).
6. Abbott, L. F. (1990) *Network: Comput. Neural Syst.* **1**, 105–122.
7. FitzHugh, R. (1961) *Biophys. J.* **1**, 445–466.
8. Nagumo, J. S., Arimoto, S. & Yoshizawa, S. (1962) *Proc. IRE* **50**, 2061–2083.
9. Horn, D. & Usher, M. (1989) *Phys. Rev. A* **40**, 1036–1044.
10. Wang, L. & Ross, J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 988–992.
11. Kaczmarek, L. K. & Levitan, I. B. (1987) *Neuromodulation* (Oxford Univ. Press, New York).
12. Marder, E., ed. (1989) *Semin. Neurosci.* **1**, 1–79.
13. Winfree, A. T. (1967) *J. Theor. Biol.* **16**, 15–42.
14. Kuramoto, Y. & Nishikawa, I. (1987) *J. Stat. Phys.* **49**, 569–605.
15. Ermentrout, G. B. & Kopell, N. (1984) *SIAM J. Math. Anal.* **15**, 215–237.
16. Hebb, D. O. (1949) *The Organization of Behavior: A Neuropsychological Theory* (Wiley, New York).
17. Kelso, S. R., Ganong, A. H. & Brown, T. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5326–5330.
18. diPrisco, G. V. (1984) *Prog. Neurobiol.* **22**, 89–102.
19. Levy, W. B. (1985) in *Synaptic Modification, Neural Selectivity and Nervous System Organization*, eds. Levy, W. B., Anderson, J. A. & Lehmkuhle, S. (Erlbaum, Hillsdale, NY).

20. Brown, T. H., Chapman, P. F., Kairiss, E. W. & Keenan, C. L. (1988) *Science* **242,** 724–728.
21. Rauschecker, J. P. & Singer, J. P. (1981) *J. Physiol. (London)* **310,** 215–239.
22. Singer, W. (1985) in *Synaptic Modification, Neural Selectivity and Nervous System Organization*, eds. Levy, W. B., Anderson, J. A. & Lehmkuhle, S. (Erlbaum, Hillsdale, NY).
23. Bear, M. F., Cooper, L. N. & Ebner, F. F. (1987) *Science* **237,** 42–48.
24. Reiter, H. O. & Stryker, M. P. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 3623–3627.
25. Gustafsson, B., Wigstrom, H., Abraham, W. C. & Huang, Y.-Y (1987) *J. Neurosci.* **7,** 774–780.
26. Stanton, P. K. & Sejnowski, T. J. (1989) *Nature (London)* **339,** 215–218.
27. Amit, D. J., Wong, K. Y. M. & Campbell, C. (1989) *J. Phys. A* **22,** 2039–2046.
28. McKay, D. M. & McCulloch, W. S. (1952) *Bull. Math. Biophys.* **14,** 127–135.
29. von der Malsburg, C. (1985) *Ber. Bunsenges Phys. Chem.* **89,** 703–705.
30. von der Malsburg, C. & Schneider, W. (1986) *Biol. Cybern.* **54,** 29–40.
31. Buhmann, J. (1989) *Phys. Rev. A* **40,** 4145–4148.
32. Li, Z. & Hopfield, J. J. (1989) *Biol. Cybern.* **61,** 379–392.
33. Kammen, D. M., Inui, M. & Koch, C. (1990) *Snowbird Conference: Neural Networks for Computing*, in press.
34. Freeman, W. J. (1978) *Elect. Clin. Neurophys.* **44,** 369–376.
35. Freeman, W. J. & van Dijk, B. W. (1987) *Brain Res.* **422,** 267–276.
36. Pöppel, E. & Logothetis, N. (1986) *Naturwissenshaften* **73,** 267–268.
37. Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M. & Reitboeck, H. J. (1988) *Biol. Cybern.* **60,** 121–130.
38. Grey, C. M. & Singer, W. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 1698–1702.
39. Gray, C. M., Kónig, P., Engel, A. K. & Singer, W. (1989) *Nature (London)* **338,** 334–337.
40. Wilson, H. R. & Cowan, J. D. (1972) *Biophys. J.* **12,** 1–24.
41. Wang, D. L., Buhmenn, J. & von der Malsburg, C. (1990) *Neural Comput.* **2,** 94–106.