

Oscillations and low firing rates in associative memory neural networks

Joachim Buhmann

Computer Science Department, University of Southern California, University Park, Los Angeles, California 90089-0782

(Received 22 May 1989)

A neural network with excitatory neurons for associative storage and inhibitory neurons for control of firing rates is proposed. In distinction to attractor neural networks, which are endowed with fixed-point dynamics, the basic recall mode of the network consists of a relaxation to a limit cycle originating from an inhibitory feedback loop. Nonlocal synaptic connections between excitatory neurons store all the information and yield robust associative abilities of the network. Inhibitory neurons with short-range connections and nonlinear interaction (shunting) are introduced to stabilize low levels of neural activity. The mean firing rate per neuron ranges between 0.1 and 0.5 impulses per Monte Carlo step (MCS). The average activity of excitatory and inhibitory neurons oscillates with a frequency of 0.5/MCS. The model generalizes the attractor concept for associative memory and brings logical neural networks closer to biological reality.

I. INTRODUCTION

Associative memories are among the earliest and most successful applications of neural networks.^{1,2} Networks of suitably interconnected simple logical units endowed with a dissipative dynamics relax to a stable network state which might be interpreted as a stored pattern. This basic concept of associative storage and recall turns out to be very robust against minor variations in the storage procedure (learning rule) and the networks dynamics. However, the concept of computing by attractors, i.e., relaxation to fixed points of network dynamics, neglects the rich temporal structure of biological neural network behavior. In this paper, two of these discrepancies, i.e., oscillations in the neural activity and low average firing rates, are addressed and an associative network is proposed which brings logical neural networks closer to their physiological counterparts.

Recent neurophysiological findings in the visual cortex³ and earlier results from the olfactory cortex⁴ suggest that oscillations in neural activity might be an important aspect of neural information processing. The precise functional role of activity oscillations in the cortex has not yet been revealed in experiments, but theoretical considerations based on neural modeling⁵ suggest that neural oscillations might be important for expressing relations between different parts or different qualities (as color, motion, texture, and depth in visual perception) of an object to be represented.

The second neurophysiological observation which stands in contrast to the dynamics of logical neural networks is the low temporal firing rate of neurons. Single-electrode recordings clearly show that cortical neurons fire between 20 and 50 impulses per second.⁶ In neural tissue preparations, however, neurons can be stimulated to fire several hundred times a second. This indicates that cortical neurons normally fire with a low rate compared to the saturation frequency set by membrane properties. Both physiological observations, oscillations and low firing rates, are not modeled realistically in attractor neural networks. Logical neural networks with associative proper-

ties relax to stable, nonoscillating states which are characterized by a subset of neurons firing with the maximal rate and all other neurons silent.

In this paper I propose a neural network with oscillating neural populations as a basic model for associative dynamics. Excitatory and inhibitory neurons⁷ play different functional roles in contrast to the Hopfield model where neurons with excitatory and inhibitory synapses are postulated. Only excitatory neurons are employed for storage of information. Excitatory interactions, which are nonlocal according to the common view in neuroanatomy,⁸ yield a cooperative dynamics with resulting fault tolerance and robustness for associative pattern storage and recall. The inhibitory neurons serve as activity control elements. They are arranged on a two-dimensional layer where each neuron interacts with neighboring excitatory and inhibitory neurons in a feedback loop. To model shunting effects of inhibitory neurons and the preference of inhibitory synapses to terminate at the cell body I have included a nonlinear term in the inhibitory interaction. The local arrangement of inhibitory cells allows an effective suppression of high activity fluctuations and emulates a fast inhibitory response. The network architecture differs from the center-surround organization of receptive fields found in visual cortex and corresponds to anatomical observations in the cerebral cortex which assert local inhibition by stellate cells.⁸ The model generalizes the concept of attractor neural networks with fixed-point dynamics and includes limit cycles and stochastic oscillations as recall modes for pattern association. The oscillations occurring in the network result from circuit properties and are not due to oscillatory behavior of individual neurons. Neural firing rates in the network are low in agreement with experimental data.

II. DEFINITION OF THE MODEL

The model neural network consists of two types of logical neurons— N excitatory and N inhibitory neurons. The internal state of excitatory (inhibitory) neurons are denot-

ed by Boolean variables $S_i^e(S_i^i)$. $S_i^{e,i}=1$ represents one action potential of neuron i , $S_i^{e,i}=0$ represents neuron i in the resting state. The patterns ξ^v stored in the network are randomly correlated N -bit words $\{1,0\}^N$. Patterns ξ^v are chosen according to the distribution

$$P(\xi_i^v) = a\delta(\xi_i^v - 1) + (1-a)\delta(\xi_i^v) \quad \forall v, i$$

where the spatial activity parameter $a \in [0,1]$ determines the percentage of 1's in ξ^v . Small values for a as chosen in the following simulations, e.g., $a \leq 0.1$, correspond to sparse coding information and yield very efficient associative storage in attractor neural networks.^{9,10}

The interaction between excitatory neurons defines the memory traces of the p patterns stored in the network. Using Hebb's postulate, all neurons belonging to one pattern fire in a correlated way and develop strong mutual connections. The Hebb rule may be mathematically defined by

$$W_{ik}^{e-e} = \frac{1}{\bar{m}aN} \sum_{v=1}^p \xi_i^v \xi_k^v. \quad (1)$$

If two neurons are both active in one pattern they are connected by a synapse of strength $W^{e-e} = 1/(a\bar{m}N)$. \bar{m} is the average temporal activity of neurons. Rule (1) forms subsets of excitatorily connected cooperating neurons. Two neurons which are not coactive in any pattern do not interact and do not develop a synapse.¹¹

Inhibitory neurons are introduced in the model to control the average temporal activity \bar{m} , i.e., to stabilize the network in a state where each neuron fires once and then stays silent for several Monte Carlo cycles (MCS). Inhibitory neurons are not engaged in associative storage. The effect of inhibition in biological networks strongly depends on the location of inhibitory synapses. Synapses, terminating at the cell body of an excitatory neuron, can play the role of a logic switch and are able to totally suppress any activity of the neuron. On the other hand, inhibitory synapses terminating in the dendritic tree of excitatory cells have a graded effect on the cell membrane potential. Since the precise synaptic geometry is not known I introduce an adjustable amount of nonlinearity in the inhibitory postsynaptic potential.

An inhibitory neuron i is locally connected and receives input from all excitatory and inhibitory neurons in its neighborhood \mathcal{N}_i of size C . Synapses from excitatory cells to inhibitory cells have the strength $W_{ik}^{i-e} = a/(a\bar{m}C)$, inhibitory cells are mutually connected by synapses of strength $W_{ik}^{i-i} = \gamma/(a\bar{m}C)$. The local field of inhibitory neurons assumes the value

$$h_i^i = \sum_{k \in \mathcal{N}_i} W_{ik}^{i-e} S_k^e - f^i \left(\sum_{k \in \mathcal{N}_i} W_{ik}^{i-i} S_k^i \right), \quad (2)$$

with

$$f^i(x) = (1 - \eta^i)x + \eta^i x^2,$$

η^i being a nonlinearity parameter for the interaction among inhibitory cells.

Inhibitory synapses ending at excitatory cells have the strength $W_{ik}^{e-i} = \beta/(a\bar{m}C)$. The resulting field h_i^e of an ex-

citatory neuron i is given by

$$h_i^e = \sum_v \frac{\xi_i^v m^v}{\bar{m}} - f^e \left(\sum_{k \in \mathcal{N}_i} W_{ik}^{e-i} S_k^i \right), \quad (3)$$

with

$$f^e(x) = (1 - \eta^e)x + \eta^e x^2.$$

η^e controls the degree of nonlinearity for inhibition of excitatory cells. The first term in (3) describes the interaction between excitatory neurons. The variable $m^v = (1/aN) \sum_i \xi_i^v S_i^e$ measures the overlap of the network of excitatory neurons with the stored pattern v . If the network is in state $S^e = \xi^v$ we obtain $m^v = 1$ and $m^\mu \approx a \forall \mu \neq v$. The second term in (3) models the nonlinear influence of inhibitory neurons on excitatory neurons.

The neurons are updated asynchronously. A randomly chosen neuron i with local field $h_i^{e,i}(t)$ fires with probability

$$P^i(t) = \{1 + \exp[-(h_i^{e,i} - U^{e,i})/T^{e,i}]\}^{-1}$$

at time $t + \tau/N$, otherwise it is silent. The parameters $U^{e,i}$ and $T^{e,i}$ are the threshold potentials and the network temperatures for excitatory and inhibitory neurons, respectively. The asynchronous update of logical neurons establishes a characteristic time scale of τ/N for one spin flip where τ corresponds to one MCS. The average update time has to be compared with the average interval between two impulses of an active cortical neuron. A neuron reaching the threshold potential needs approximately 1 ms for the action potential, 3–5 ms for the absolute refractory period, 2–5 ms for the relative refractory period, and additional 1–3 ms for integration of inputs until it reaches the threshold value again. In total, we have about 7–14 ms as the equivalent physiological time for one MCS depending on the cell type and the local neuroanatomy and neurochemistry.

The network composed of excitatory and inhibitory neurons has a much more complex connectivity pattern than the standard model.² The connections between excitatory and inhibitory cells are nonsymmetric; dynamics of the system is not dominated by an energy function, i.e., we can expect stochastic oscillations.

III. "FERROMAGNETIC" NETWORK WITH LOCAL INHIBITION

Before we discuss the behavior of the network with several patterns stored we will study the much simpler case of a network with complete connectivity between excitatory neurons. One can consider such a system as a neural network with only one pattern stored. The network is formally equivalent to a long-range ferromagnet with local inhibition.

To monitor the network behavior and to measure the average temporal activity of a neuron, we introduce the variables $x^{e,i} = (1/aN) \sum_i \langle S_i^{e,i} \rangle$. $x^{e,i}$ measure the average activity of excitatory and inhibitory neurons ($\langle \dots \rangle$ denotes temporal averaging). These values have to be compared with the variables $q^{e,i} = (1/aN) \sum_i \langle S_i^{e,i} \rangle^2$, which measure the degree of freezing. Low average temporal

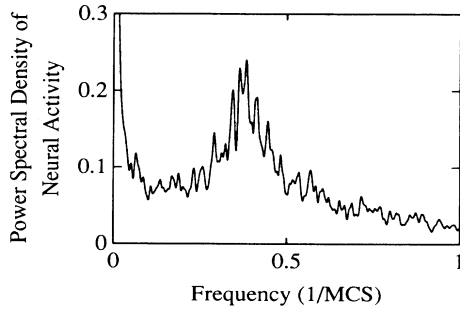


FIG. 1. Power spectral density of the excitatory activity in a ferromagnet with local inhibition ($N=1000$, $C=81$, $\alpha=1$, $\beta=1.5$, $\gamma=1$, $U^e=0.2$, $U^i=0.6$, $T^e=T^i=0.1$, $\eta^e=\eta^i=0$).

firing rate of all neurons, the desired behavior of the network, is indicated by $q^{e,i} \approx (x^{e,i})^2$. The opposite case $x^{e,i} \approx q^{e,i}$ stands for the high firing rates of a few neurons ($x^{e,i}$ percent) and no activity of the rest. This case corresponds to low spatial activity networks.

The global behavior of the network depends on the synaptic values, the thresholds and the noise levels. For very high temperature ($T^{e,i} \gg U^{e,i}$) the network is in the paramagnetic state. For vanishing noise ($T^{e,i}=0$) and positive thresholds $U^{e,i}$, the network state with zero activity $x^{e,i}=0$ is always a stable fixed point of the dynamics. A richer dynamical behavior than fixed-point relaxation can be observed if the initial activity of excitatory neurons exceeds a lower bound ($m^e > U^e \bar{m}$) and if the noise level is reduced below a critical transition temperature. Monte Carlo simulations of a ferromagnetic net with local inhibition ($N=1000$) show oscillations of the average activity of neurons. After 100 MCS the values $x^e=0.186$, $q^e - (x^e)^2 = 4 \times 10^{-4}$, $x^i=0.132$, and $q^i - (x^i)^2 = 8 \times 10^{-5}$ clearly indicate low firing rates of excitatory and inhibitory neurons. The power spectral density of activity oscillations is shown in Fig. 1. Stable oscillations in the system are identified by a maximum around 0.4/MCS. If we interpret a Monte Carlo step as 10 ms we get 40-Hz oscillations, as observed recently in visual cortex³ and in olfactory bulb.⁴ The oscillations in the ferromagnet with local inhibition are quite stable for different parameter values. The oscillation frequency varies around 0.3–0.6/MCS and depends on the locality of inhibition, e.g., for nonlocal inhibition the excitatory pool oscillates with 0.7/MCS.

The nonlinearity (η^e) in the inhibition of excitatory neurons increases the stability of the low-activity network state. The nonlinearity, however, disturbs the regular oscillations. The fairly localized spectral density of a net with linear ($\eta^e=0$) interaction is transformed into a broad frequency band indicating that many competitive oscillations occur in the network dynamics.

IV. ASSOCIATIVE MEMORY WITH LOCAL INHIBITION

Let us now discuss the general case of many patterns stored in the network. What does pattern recall mean in a network with oscillating neural activity? In attractor

neural networks, stable fixed points of the dynamics are interpreted as the memory output. Recall errors are identified as active neurons that are supposed to be silent or as silent neurons that are supposed to be active. In the present case of activity oscillations and low firing rates I will identify an average impulse rate of a neuron exceeding a certain threshold rate as a 1 bit and a very low or vanishing impulse rate as a 0 bit. Recall errors occur due to broadening effects of the distribution of impulse rates, i.e., neurons active in a pattern state might fire too seldom and, therefore, are classified as silent. The opposite case of activating a neuron which should be silent is also a possible source of recall errors.

The recall procedure in associative neural networks can be defined in two different ways: (i) a disturbed or incomplete pattern is chosen as the initial network state and the cooperative dynamics of the network provides the missing information; (ii) a disturbed or incomplete pattern is presented as constant input to the network over a sufficiently long period (≈ 5 MCS) and the network evolves to a state or a sequence of states which correspond to the complete pattern. I have successfully tested the oscillating network in both recall modes. Simulation results of the first mode are reported in this section; simulations with a network operating in the second mode will be discussed elsewhere.

A typical evolution of the overlap m^1 between the network state and pattern 1 (solid line), the background activity (dotted line), and the inhibitory activity (dashed line) is shown in Fig. 2. The network ($N=2000$, $p=20$) was prepared in the initial state S_i^0 , characterized by probability $P(S_i(0)=1) = \bar{m}\xi_i^1 + 0.5\bar{m}\xi_i^2$. After 4 MCS the network relaxed to a phase-space region with $m^1 \in [0.15, 0.3]$, $m^\mu = a\bar{m}\forall \mu > 1$. The initial admixture of the second pattern was quickly suppressed and the first pattern restored. The distribution of firing rates of active excitatory neurons ($\xi_i^1=1$) is peaked around 0.2/MCS with a width of 0.1/MCS indicating that a neuron fires approximately once every 5 MCS. Only very few background neurons ($\xi_i^1=0$) fire with rates exceeding the value 0.025/MCS and cause recall errors. If associative recall of a stored pattern fails, the network settles down in the state of zero activity $x^{e,i}=0$ which serves as an indica-

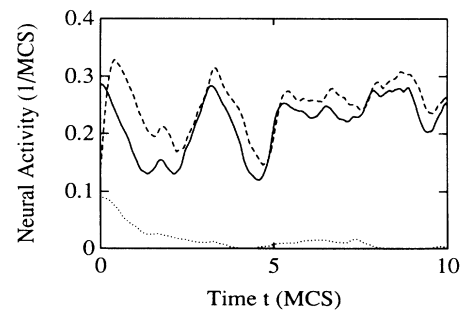


FIG. 2. Evolution of overlap m^1 (solid), inhibitory activity $\sum_k S_k^i/aN$ (dashed), and background activity $\sum_k S_k^0(1-\xi_k^1)/aN$ (dotted) during associative recall of pattern 1 ($N=2000$, $p=20$, $C=361$, $a=0.1$, $\bar{m}=0.2$, $\alpha=\gamma=1$, $\beta=0.9$, $U^e=0.2$, $U^i=0.6$, $T^e=T^i=0.05$, $\eta^e=0.25$, $\eta^i=0$).

tor of incomplete association.

The question of storage capacity in oscillating networks, i.e., the ratio of stored patterns p to excitatory neurons N , is subject of current research. A slight modification of the excitatory interaction (1) in the case $p \sim N$ is necessary to bias the mutual overlap of different patterns (structural noise). The mean value of the structural noise can be shifted to 0 by introducing inhibitory neurons with a global connectivity. I expect a qualitatively comparable network efficiency, as for attractor neural networks with low spatial activity.

V. CONCLUSION

The proposed model neural network generalizes the principle of fixed-point relaxation used for associative storage and recall in attractor neural networks. Oscillations around a network state of low average neural activity are interpreted as the network response to a stable input pattern corrupted by noise or to an initial state close to one stored pattern. Modeling the different functional

roles of excitatory and inhibitory neurons, the network architecture gives rise to oscillations of the average neural activity, as are actually observed in the visual and olfactory cortex. These oscillations originate from the asymmetric coupling between populations of excitatory and inhibitory neurons. The network also produces low firing rates for all neurons active in a pattern state. Both properties, oscillations and low firing rates, bring the network closer to biological observation and help to close the gap between biological neural networks and their Boolean counterparts.

ACKNOWLEDGMENTS

The author is grateful to C. von der Malsburg for stimulating discussions on oscillations in neural systems and to D. A. Amit and A. Treves for pointing out the *high firing rate* problem in attractor neural networks. This work was supported by a NATO Grant No. DAAD 300/402/513/9 and by U.S. Air Force Office of Scientific Research Grant No. 88-0274.

¹K. Steinbuch, *Kybernetik* **1**, 36 (1961); D. J. Willshaw, O. P. Bunemann, and H. C. Longuet-Higgins, *Nature* (London) **222**, 960 (1969).

²J. J. Hopfield, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554 (1982).

³C. M. Gray, P. König, A. K. Engel, and W. Singer, *Nature* (London) **338**, 334 (1989); R. Eckhorn, R. Bauer, W. Jordan, M. Brosch, W. Kruse, M. Munck, and H. J. Reitboeck, *Biol. Cybernet.* **60**, 121 (1988).

⁴W. J. Freeman, *Electroenc. Clin. Neurophys.* **44**, 586 (1977).

⁵C. von der Malsburg and W. Schneider, *Biol. Cybernet.* **54**, 29 (1986).

⁶M. Abeles, *Local Cortical Circuits* (Springer, New York, 1982); Y. Miyashita, *Nature* (London) **335**, 817 (1988).

⁷Independent from my studies D. A. Amit and A. Treves have

proposed an associative memory network with fast inhibition to solve the high-firing-rate problem; D. J. Amit and A. Treves, *J. Phys. A* **22**, 2205 (1989).

⁸E. R. Kandel and J. R. Schwartz, *Principles of Neural Science* (Elsevier, Amsterdam, 1985), p. 636.

⁹J. Buhmann, R. Divko, and K. Schulten, *Phys. Rev. A* **39**, 2689 (1989).

¹⁰M. V. Tsodyks and M. V. Feigel'man, *Europhys. Lett.* **6**, 101 (1988).

¹¹To economize memory space in the simulations, I have restricted W_{ik}^c to a synaptic depth of 1 bit, i.e., $W_{ik}^c = 1/(\bar{m}aN)$, if there exists a pattern v with $\xi_i^v \xi_k^v = 1$ and $W_{ik}^c = 0$ otherwise.