Chaos, multiplicity, crisis, and synchronicity in higher-order neural networks

Lipo Wang and John Ross

Department of Chemistry, Stanford University, Stanford, California 94305
(Received 1 July 1991)

We study a randomly diluted higher-order network of spinlike neurons that interact via Hebbian-type connections and derive and solve exact dynamical equations for a general block-sequential updating algorithm. The system has a variety of static and oscillatory solutions. The bifurcation parameters in the present model include neuronal interaction coefficients, the synchronicity parameter, and a re-scaled noise level, which represents the combined effects of the random synaptic dilution, interference between stored patterns, and additional background noise.

Recently there have been extensive research activities in the area of artificial neural networks (NN's) [1–11]. Model neurons have two states, i.e., active and quiescent. Each neuron receives signals from its neighboring neurons, and the signals are affected by synaptic weights. The neuron then either fires if the total signal exceeds a certain threshold, or remains quiescent otherwise. These spinlike systems, though crude compared to biological NN's, already display intriguing features, such as a form of learning and recalling of associative memory. When the synapses, which are analogous to spin-spin interactions, are symmetric and the states of neurons are updated sequentially, the NN can be analyzed by statistical mechanics [4], which has provided important information on the role of noise, memory capacity, and retrieval performance. The network dynamics for sequential updating is a simple nonoscillatory relaxation to an energy minimum [3]. For symmetric interactions but parallel updating, i.e., all neurons are updated simultaneously at one time, the network settles down to either a stable state or a cycle of period 2 (alternating between two states) [5]. When the synapses are asymmetric, as in real neural systems, the dynamics can be oscillatory or chaotic, e.g., in a NN with *random synapses* [10]; however, no patterns can be stored in this NN with *random synapses* since no learning rules are used.

In the present paper we study a diluted [9] NN model that uses a higher-order modified Hebb learning rule [3,6,12,13] so that patterns can be stored in this system. Our model exhibits a variety of dynamical behaviors besides stable retrieving, such as oscillations, chaos, multiplicity, and crisis. In this approach, we consider N two-state neurons that interact through q orders of Hebbian-type connections. Thus the total input for the ith neuron is

$$h_i(t) = \sum_{v=1}^{q} \sum_{j^{(1)},\ldots,j^{(v)}=1}^{N} T_{ij^{(1)}\ldots j^{(v)}} S_{j^{(1)}}(t) \cdots S_{j^{(v)}}(t) + \eta_i ,$$

(1)

where $S_j(t) = \pm 1$ represents the state of the neuron j at time t,

$$T_{ij^{(1)}\ldots j^{(v)}} = C_{ij^{(1)}\ldots j^{(v)}} \gamma^{(v)} \sum_{\mu=1}^{p} S_i^\mu S_{j^{(1)}}^\mu \cdots S_{j^{(v)}}^\mu$$

(2)

are the modified Hebbian synaptic efficacies, $\{S_j^\mu, j$

$=1,\ldots,N\}$ is the μth stored pattern which is an N-bit word and each bit is ± 1 with equal probability, and p is the number of patterns stored. The coefficients $\{\gamma^{(v)}, v=1,\ldots,q\}$ measure the relative strengths of different order of interactions [6]. We have introduced random *asymmetric* dilution in the effiacies by choosing random variables $C_{ij^{(1)}\ldots j^{(v)}}$ *independently* according to the following distributions:

$$\rho(C_{ij^{(1)}\ldots j^{(v)}}) = \left[C \Big/ \binom{N}{v}\right] \delta(C_{ij^{(1)}\ldots j^{(v)}} - 1)$$
$$+ \left[1 - C \Big/ \binom{N}{v}\right] \delta(C_{ij^{(1)}\ldots j^{(v)}}) .$$

(3)

With $v=1$, (3) reduces to a prescription first used by Derrida, Gardner, and Zippelius (DGZ) in their discussion of a first-order diluted network [9]. Synaptic dilution is essential in both modeling the observed incomplete connectivity in real neurophysiological systems and assuring an *exact* solution of the network dynamics [9]. We include a background random *Gaussian* [14] noise η_i with a standard deviation σ_0 in (1) to account for the presence of noise (temperature) [2].

The neuron i updates its state according to $S_i(t+\Delta t) = \text{sgn}[h_i(t)]$, where $\text{sgn}(x) = -1 \ (+1)$ for negative (positive) x. Suppose that the initial state of the network is set in the neighborhood of stored pattern $\mathbf{S}^1$, i.e., $m^1(0) = \max\{m^\mu(0), \ \mu=1,2,\ldots,p\}$, where $m^\mu(t) = (1/N)\mathbf{S}^\mu \cdot \mathbf{S}(t)$ is the overlap between the state of the system at time t and the μth pattern.

Let $\{j_1^{(1)} \cdots j_1^{(v)}\}, \ldots, \{j_{K^{(v)}}^{(1)} \cdots j_{K^{(v)}}^{(v)}\}$ be the $K^{(v)}$ groups of neurons $\{j^{(1)} \cdots j^{(v)}\}$ such that $T_{ij^{(1)}\ldots j^{(v)}} \neq 0$. According to (3), for all $v=1,\ldots,q$, the average of $K^{(v)}$ is C, compared with fully connected networks where the number of synaptic connections of order v is $\binom{N}{v}$.

We substitute (2) into (1) and separate the first term in the total local field, i.e., the sum in (1), into two parts:

$$h_i(t) = S_i^1 \sum_{v=1}^{q} \gamma^{(v)} \sum_{r=1}^{K^{(v)}} m_{j_r^{(1)}}(t) \cdots m_{j_r^{(v)}}(t) + \zeta_i + \eta_i ,$$

(4)

where $m_j(t) = S_j^1 S_j(t)$. The first term in (4) is proportional to the overlap of the system with pattern $\mathbf{S}^1$. We have dropped the superscript 1 for the overlaps and hereafter we always refer to the overlap with $\mathbf{S}^1$. The second term ζ_i is the residual and consists of interferences from patterns $\mathbf{S}^2,\ldots,\mathbf{S}^p$, which are *independent* random vari-

ables with mean of zero if the neuronal states involved in (4) are *uncorrelated*. As shown in Ref. [9] for the case of first-order interaction $(q=1)$, these correlations can be neglected in the limit of extreme dilution. We now prove that this result holds in the presence of higher-order interactions as given by (1)–(3). The calculation of the state of the ith neuron at time t, i.e., $S_i(t)$, involves a tree of ancestors (states of neurons at previous time steps) which connects the ith neuron to the initial conditions $(t=0)$. At each time step the state of the ith neuron is influenced by about vC neurons through interactions of order v, with $v=1,\ldots,q$, so the total number of neurons in this tree is typically of the order of $[q(q+1)C/2]^t$. Thus as long as $C \ll \ln N$, almost all neurons in this tree are different. Therefore, in this diluted limit, the neuronal states involved in (4) are uncorrelated. According to the central limit theorem, the interference ζ_i is random and Gaussian distributed, with a total average of zero and a total squared deviation

$$\sigma_{\text{ct}}^2 \equiv (p-1) \sum_{v=1}^{q} [\gamma^{(v)}]^2 K^{(v)}$$

in the limit of large N. From this expression of σ_{ct}, we see that no generality has been lost by choosing the same dilution constant C for all orders of interactions in (3), since the effects of different dilution constants can be absorbed into coefficients $\gamma^{(v)}$. The combination of the interference ζ_i and the background Gaussian noise η_i gives a new Gaussian random variable with mean of zero and a squared deviation $\sigma_{\text{total}}^2 = \sigma_{ct}^2 + \sigma_0^2$. Furthermore, when we average over all shapes of the tree of ancestors, all the $m_j(t)$ have the same average $m(t)$ [9]. It follows from (4) that

$$m_i(t+1) = \text{erf}\left[\frac{1}{\sqrt{2}\sigma} \sum_{v=1}^{q} \gamma^{(v)}[m(t)]^v\right], \qquad (5)$$

with a rescaled squared standard deviation of the noise

$$\sigma^2 \equiv [(p-1)/C] \sum_{v=1}^{q} [\gamma^{(v)}]^2 + (\sigma_0/C)^2,$$

which includes the effects of random synaptic dilution, interference between stored patterns, and background noise.

We use a general block-sequential updating algorithm [5] and select U neurons to update. We have from (5)

$$m(t+\Delta t) = (1-u)m(t) + u\,\text{erf}\left[\frac{1}{\sqrt{2}\sigma} \sum_{v=1}^{q} \gamma^{(v)}[m(t)]^v\right]$$

$$\equiv f_u^{(q)}(m(t),\sigma), \qquad (6)$$

where $u = U/N$ is the synchronicity parameter. Parallel dynamics corresponds to $u=1$. In sequential dynamics $u = \Delta t = 1/N \to 0$, as $N \to \infty$, and (6) reduces to a differential equation:

$$\frac{dm(t)}{dt} = f_0^{(q)}(m(t),\sigma) - m(t). \qquad (7)$$

Equations (6) and (7) are a generalization of the formulation by DGZ [9], who derived an exact solution of a diluted NN with first-order interactions $(q=1)$ where the dynamical equations for the overlap $m(t)$ are nonoscilla-

tory. Only fully parallel $(u=1)$ and sequential updating $(u \to 0)$ were considered [9]. Furthermore, they used Little's definition of temperature [2] instead of the more biologically plausible Gaussian noise used in the present work.

In the following, we show that the overlap $m(t)$ in the present model exhibits a variety of interesting dynamical phenomena, such as oscillations, period-doubling bifurcations, chaos, multiplicity, crisis, as well as stable retrievals, *if higher-order interactions exist* $(q > 1)$. We also discuss how synchronicity in updating $(0 < u \le 1)$ influences the network dynamics.

Let us first consider sequential updating $(u \to 0)$, where the network dynamics is described by (7). A one-dimensional system described by a differential equation cannot exhibit oscillatory behavior [15]. Hence sequential dynamics has only two types of *nonoscillatory* fixed points m^*, which can be either stable or unstable, for any orders of interaction and coefficients $\gamma^{(v)}$. The fixed points of the network are obtained by letting $dm/dt = 0$ and $m(t) = m^*$ in (7) or $m(t+\Delta t) = m(t) = m*$ in (6). It is apparent that the fixed points are independent of synchronicity u; however, as we will show shortly, their stability depends strongly on the synchronicity. As an example, the positive and negative fixed points m^* for $q=3$, $\gamma^{(1)}=1$, $\gamma^{(2)}=-4$, and $\gamma^{(3)}=4$, are given in Figs. 1(a) and 1(b), respectively, as a function of the rescaled noise level σ. For sequential dynamics, the unstable fixed points represented by the dashed lines are the boundaries of basins of attractions of the stable fixed points (solid lines).

For any finite synchronicity $(0 < u \le 1)$, oscillatory behaviors begin to emerge as a result of some of the fixed points between C and B in Fig. 1(a) becoming unstable, whereas stabilities of the rest of the fixed points remain the same as in the sequential updating. The equation of motion of the neural network is now given by (6), which is
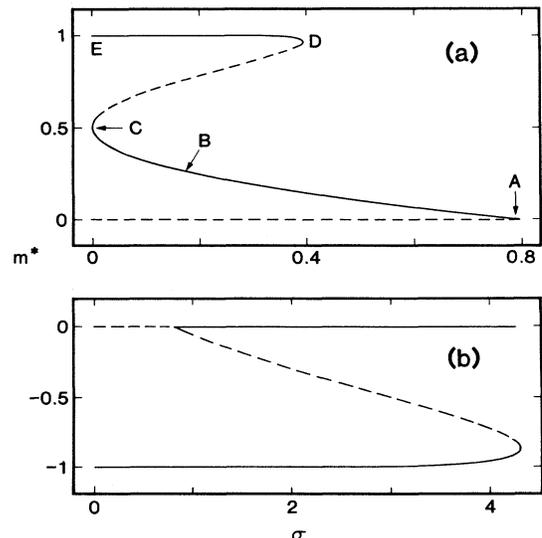
FIG. 1. Fixed points m^*, positive in (a), negative in (b), of the overlap given by (7) vs the rescaled noise σ, for $q=3$, $\gamma^{(1)}=1$, $\gamma^{(2)}=-4$, and $\gamma^{(3)}=4$.

a one-dimensional map [15] of the overlap $m(t)$, with the following bifurcation parameters: interaction coefficients $\{\gamma^{(v)}, v=1,\ldots,q\}$ and the rescaled noise level σ. For $q=3$, $\gamma^{(1)}=1$, $\gamma^{(2)}=-4$, and $\gamma^{(3)}=4$, the mapping function $f_u^{(3)}(m)$ has, depending on σ, up to four intersecting points with the diagonal line within the invariant box $0 \le m \le 1$, $0 \le f_u^{(3)}(m) \le 1$, and these are the fixed points. The smallest nonzero intersecting point is responsible for the oscillatory dynamics. Since $0 \le \mathrm{erf}(x) \le 1$, for $x \ge 0$, we have

$$(1-u)m \le f_u^{(3)}(m) \le (1-u)m+u .$$

Hence amplitudes of oscillations increase (actual bifurcation diagrams omitted) with increasing synchronicity u and reach maxima in fully parallel updating algorithm $(u=1)$, for which we show the mapping function at various σ in Fig. 2. For $\sigma > \sigma_1 \equiv 0.191$, the system displays fixed point dynamics (curve a). As the noise level decreases below σ_1, the slope at the smallest nonzero intersecting point becomes less than -1 and oscillations start to appear (curve b in Fig. 2). As shown in Fig. 3, there is a complete period-doubling sequence between σ_1 and $\sigma_\infty \equiv 0.142$. Chaotic structures and periodic windows can be seen beyond the saturation point σ_∞, e.g., a period-3 window is marked in Fig. 3. For $\sigma_1 > \sigma > \sigma_{\mathrm{crisis}} \equiv 0.076$, the oscillatory structures coexist with the upper stable fixed points [marked D in Fig. 2, portion ED in Fig. 1(a)]; the second smallest intersecting points, e.g., point B and C in Fig. 2, serve as a separatrix [portion CD in Fig. 1(a)]. As σ is decreased below σ_{crisis}, oscillatory and chaotic structures suddenly disappear at σ_{crisis}, which is called "crisis" (marked C in Fig. 3). In Fig. 2 the portion of curve c represented by a solid line is higher than its corresponding separatrix C and hence acts as a "gate" to the upper stable fixed points D: all iterations started between $m(0)=0$ and the separatrix C will go through this gate and will then be attracted to D. When σ is below σ_{crisis} by
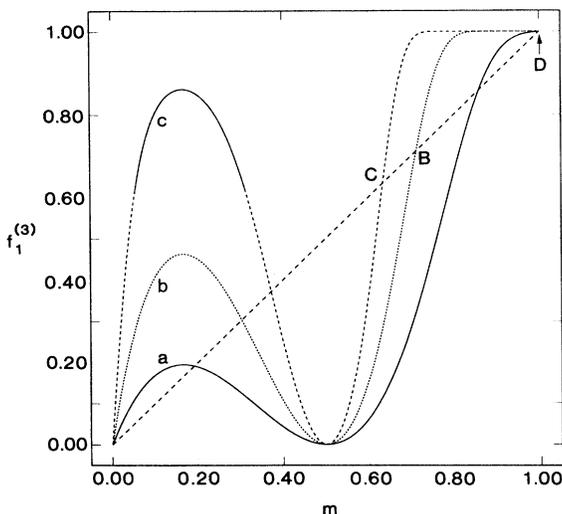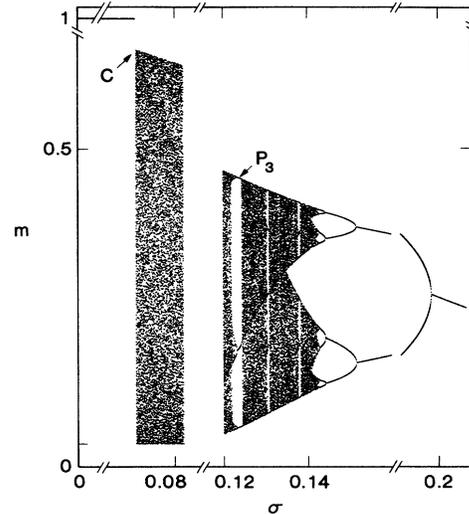
FIG. 3. Oscillatory and chaotic structures for parallel dynamics $u=1$.

only a small amount, it takes a long time for the network to "find" the small gate and then the attractor, which is referred to as transient chaos.

The above discussions based on analytical results and numerical simulations are subject to future work. We expect to see agreement between our results and simulations if implemented networks are (exponentially) large. We have concentrated our discussions for oscillations and chaos $(u>0)$ on some special choices of q, and $\gamma^{(v)}$. Discussions for an arbitrary case are similar because the dynamical equation (6) is always a one-dimensional map and dynamical universalities of one-dimensional maps have been thoroughly explored [15]. For instance, if $f_u^{(q)}$ has one or more maxima, i.e., $q \ge 2$, period-doubling bifurcations become possible; for $q \ge 3$, $f_u^{(q)}$ may have two or more maxima and thus crisis may occur, though actual bifurcation diagrams will differ quantitatively from case to case.

We have presented a simple NN model that is exactly solvable and that exhibits interesting phenomena such as oscillations, chaos, multiplicity, and crisis. These dynamics are the result of an interplay among the synchronicity, the asymmetric dilution, and higher order interactions. Oscillatory behaviors are not likely for *symmetric* networks with *higher-order interactions*, since we can generalize the Hopfield [3] energy function for symmetric first-order NN to any arbitrary orders of *symmetric* interactions:

$$E^{(q)} = -\sum_{v=1}^{q} 1/(v+1) \sum_{i,j^{(1)},\ldots,j^{(v)}=1}^{N} T_{ij^{(1)}\ldots j^{(v)}} S_i(t)$$
$$\times S_{j^{(1)}}(t) \cdots S_{j^{(v)}}(t) ,$$

where by "symmetric" we mean that all $T_{ij^{(1)}\ldots j^{(v)}}$ remain unchanged when any two subscripts are exchanged.

Oscillations and chaos have been discussed in relation to perception processes by a number of authors [16]. The discovery of 40-Hz oscillations in electroencephalograms

FIG. 2. The mapping function $f_1^{(3)}(m)$ given by (6) vs m: a, $\sigma=0.3$ (stable fixed point); b, $\sigma=0.12$ (oscillation); c, $\sigma=0.05$ (crisis).

seems to support a new point of view that biological NN's use chaotic attractors, rather than fixed-point attractors as suggested in conventional NN theories, to store and process memory [16]. Cyclic and chaotic attractors around reference patterns (S^μ) may be stored in and recalled by using our prescriptions; in earlier work temporal sequences were usually generated by means of delays [7,8].

---

[1] For a recent review on NN's, see H. Sompolinsky, Phys. Today **41** (12), 70 (1988).

[2] W. A. Little, Math. Biosci. **19**, 101 (1974).

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

[4] D. J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. Lett. **55**, 1530 (1985).

[5] E. Goles-Chacc, F. Fogelman-Soulie, and D. Pellegrin, Discrete Appl. Math. **12**, 261 (1985).

[6] P. Peretto and J. J. Niez, Biol. Cybern. **54**, 53 (1986).

[7] H. Sompolinsky and I. Kanter, Phys. Rev. Lett. **57**, 2861 (1986).

[8] D. Kleinfeld, Proc. Natl. Acad. Sci. U.S.A. **83**, 9469 (1986).

[9] B. Derrida, E. Gardner, and A. Zippelius, Europhys. Lett. **4**, 167 (1987).

[10] H. Sompolinsky, A. Crisanti, and H. J. Sommers, Phys. Rev. Lett. **61**, 259 (1988).

[11] L. Wang and J. Ross, Proc. Natl. Acad. Sci. U.S.A. **87**, 988 (1990); **87**, 7110 (1990).

[12] D. O. Hebb, *The Organization of Behavior* (Wiley, New York, 1949), p. 44.

[13] G. L. Shaw, E. Harth, and A. B. Scheibel, Exp. Neurol. **77**, 324 (1982).

[14] M. Abeles, *Local Cortical Circuits* (Springer-Verlag, New York, 1982), p. 21.

[15] J. M. T. Thompson and H. B. Stewart, *Nonlinear Dynamics and Chaos* (Wiley, New York, 1986).

[16] W. J. Freeman, Sci. Am. **264** (2), 78 (1991).