Asymmetric Neural Networks Incorporating the Dale Hypothesis and Noise-Driven Chaos

T. Fukai[1] and M. Shiino[2]

[1]*Department of Management and Information Science, Gumma Women's College, Naka-Oorui-cho 501, Takasaki, Japan*
[2]*Department of Applied Physics, Tokyo Institute of Technology, Ohokayama, Meguro-ku, Tokyo, Japan*

(Received 15 August 1989)

Dynamical properties of the neural networks with asymmetrical synaptic couplings respecting the Dale hypothesis are studied. The time evolution of the networks is assumed to obey stochastic dynamics of the Little type with time delay. Using a nonlinear master equation, exact equations are derived for the time evolution of the overlaps of instantaneous configuration with p embedded patterns and with the characteristic pattern representing the configuration of excitatory and inhibitory neurons. It is shown that the networks exhibit noise-driven chaotic motions in the retrieval process.

PACS numbers: 87.10.+e, 02.50.+s, 05.45.+b, 89.70.+c

Studying asymmetric neural networks[1-3] is of biological importance since the synaptic couplings are in general asymmetric in physiological nervous systems like the brain, while we have only a little knowledge of the rules which govern the formation of the synaptic couplings in physiological nervous systems. Among such rules, it is generally known as the Dale hypothesis[4] that in physiological nervous systems excitation and inhibition are assigned to different sets of neurons, that is, each neuron has a unique excitatory or inhibitory character. It will be important to examine the functions of the rule in the global network performances of physiological nervous systems.[5,6]

In this paper, we study the retrieval dynamics of the model neural networks with asymmetric couplings which incorporate the Dale hypothesis in the scheme of the Hebb learning rule.[6] In connection with the introduction of the Dale hypothesis, we incorporate another neurophysiological aspect, i.e., the time delay, into the modeling of the network dynamics which we assume to take the form of Little's model[7] with a finite number of embedded patterns. We have found an interesting result that the present simple model has parameter regions in which it exhibits chaos.[8-11] Our model and the analysis yield a theoretically rigorous way to study complicated motions like chaos in the neural networks with the physiological constraint.

Appearance of the chaotic behavior was also observed experimentally in physiological nervous systems. Chaotic behavior was first found in the responses of action potential of a single neuron under periodic stimuli.[12] Furthermore, recently Freeman observed chaotic responses of action potential for the olfactory bulb of a rabbit when the unreinforced odor is given as a stimulus.[13] He suggested that the chaotic responses play essential roles while the learning of new odor takes place. At this stage of the neurophysiology, it will be useful to present a simple neural-network model based on physiological constraints and to show the appearance of chaotic memory retrieval through theoretically rigorous treatment.

We consider the following learning rule of patterns $\{\xi_i^{(\mu)}=\pm 1\}$ $(i=1,\ldots,N,\ \mu=1,\ldots,p)$ to generate the synaptic couplings obeying the Dale hypothesis:[6]

$$J_{ij}=\frac{1}{N}\sum_{\mu,\nu=1}^{p}A_{\mu\nu}\epsilon_{ij}^{\mu\nu}\xi_i^{(\mu)}\xi_j^{(\nu)},\quad \epsilon_{ij}^{\mu\nu}=1+\eta_j\xi_i^{(\mu)}\xi_j^{(\nu)},\quad (1)$$

$$i,j=1,\ldots,N,$$

where $A_{\mu\nu}$ is a real $p\times p$ matrix satisfying $A_{\mu\nu}\geq 0$ and η_j is $+1$ (-1) if the jth neuron is excitatory (inhibitory). We call $\{\eta_j\}$ the characteristic pattern hereafter since it represents the pattern of the configuration of excitatory and inhibitory neurons in the neural networks. Since $\epsilon_{ij}^{\mu\nu}$ vanishes if $\xi_i^{(\mu)}\xi_j^{(\nu)}$ has an opposite sign to η_j, the learning of the patterns by the synaptic couplings takes place only when $\xi_i^{(\mu)}\xi_j^{(\nu)}$ has an allowed sign by the excitatory or the inhibitory character of the jth neuron.[6] Thus the couplings given by Eq. (1) respect the Dale hypothesis. In general, the couplings become asymmetric even if the matrix $A_{\mu\nu}$ is symmetric. However, the degree of asymmetry is small in such cases and we can introduce larger asymmetry by choosing an asymmetric matrix for $A_{\mu\nu}$.

The couplings defined in Eq. (1) can be written in the following simple form if we substitute the expression of $\epsilon_{ij}^{\mu\nu}$ into the definition of J_{ij}: $J_{ij}=(1/N)\sum_{\mu\nu}A_{\mu\nu}\xi_i^{(\mu)}\xi_j^{(\nu)}+(1/N)\eta_j\sum_{\mu\nu}A_{\mu\nu}$. Note that the couplings depend linearly on the pattern $\{\eta_j\}$ as well as $\{\xi_j^{(\nu)}\}$. As we will see later, this linearity of the couplings enables us to derive a self-consistent mapping for $p+1$ pattern overlaps from a nonlinear master equation for 2^{p+1} probabilities[14] [see Eqs. (4) and (5)].

In designing the dynamics, which keeps the relation to the physiological nervous systems, we take account of the time delay[15] originating from the synaptic delays and/or the slow components of response in the definition of the local field $h_i(t)$ representing the total synaptic input to the ith neuron at time t. Then $h_i(t)$ amounts to a weighted average over the histories of the neural activities and is given by $h_i(t)=\sum_j J_{ij}V_j(t)$. Here, with

$\{S_i(=\pm 1)\}$ representing the macroscopic state of the neural network, $V_j(t)=\sum_{u=0}^{\infty}K(u)S_j(t-u)$ is the time-averaged output of the jth neuron with a suitable weight $K(u)\geq 0$.

We now consider the network dynamics to be a non-Markovian stochastic dynamics which formally assumes the form of Little's model.[7] To be specific, the dynamics is described by the conditional probability of finding the ith neuron in state σ at time $t+1$, given that microscopic states of the system at time τ ($\leq t$) are $\{S_i(\tau)\}$:

$$P(S_i(t+1))=\sigma\,|\,\{S_i(\tau)\},\tau\leq t)$$
$$=\tfrac{1}{2}\{1+\sigma\tanh[\beta h_i(t)]\}\,. \qquad (2)$$

In the present Letter, we deal with the simplest case for the time delay and set $K(u)=\delta_{u,0}+k\delta_{u,1}$ ($k\geq 0$). The parameter β represents a measure of inverse magnitude of external noise rather than the inverse physical temperature of the neural networks. It is, however, referred to as "inverse temperature" of neural networks throughout this paper in analogy with thermodynamical spin systems.

From the expression of J_{ij}, the effective local fields turn out to be described in terms of the overlaps $m^{(0)}\equiv(1/N)\sum_i \eta_i S_i$ of microscopic states at time t and $t-1$ with the characteristic pattern $\{\eta_i\}$ representing the configuration of excitatory and inhibitory neurons, as well as the overlaps $m^{(\mu)}\equiv(1/N)\sum_i \xi_i^{(\mu)}S_i$ with the embedded patterns $\{\xi_i^{(\mu)}\}$:

$$h_i(t)=\sum_{\mu\nu}A_{\mu\nu}\xi_i^{(\mu)}[m^{(\nu)}(t)+km^{(\nu)}(t-1)]$$
$$+\sum_{\mu\nu}A_{\mu\nu}[m^{(0)}(t)+km^{(0)}(t-1)]\,. \qquad (3)$$

We may say that specifying excitatory or inhibitory character for each neuron can be interpreted as embedding an extra pattern $\{\eta_i\}$ in the present neural networks. Accordingly the macroscopic dynamics of the pattern overlaps are now described in a $(p+1)$-dimensional configuration space rather than in a p-dimensional one.

Let us formulate explicitly the mapping describing the motions of the pattern overlaps. Note that in the discussion below the embedded patterns need not be random patterns but can be an arbitrary set of patterns. We first divide the system of N neurons into at most 2^{p+1} sublattices $\Omega(\mathbf{X})$ (Refs. 3 and 14) according to the p embedded patterns and the characteristic pattern:

$$\Omega(\mathbf{X})=\{i\,|\,\xi_i^{(1)}=\xi^{(1)},\ldots,\xi_i^{(p)}=\xi^{(p)},\eta_i=\eta\}\,,$$

where the $p+1$ vector $\mathbf{X}=(\xi^{(1)},\ldots,\xi^{(p)},\eta)$ is defined in the $(p+1)$-dimensional hypercube $\mathbf{H}^{p+1}$ of ± 1 coordinates. Then, we define the rate of appearance $r(\mathbf{X})$ of the vector $\mathbf{X}$ as $r(\mathbf{X})=\lim_{N\to\infty}|\,\Omega(\mathbf{X})\,|/N$, with $|\,\Omega(\mathbf{X})\,|$ denoting the number of elements of $\Omega(\mathbf{X})$.

Next, noting that the law of large number holds in the present system, we define, in the thermodynamical limit $N\to\infty$, the empirical probability $p(t;\mathbf{X})$ of finding $+1$ state at time t for the neurons in the sublattice $\Omega(\mathbf{X})$ as the large-N limit of the ratio to $|\,\Omega(\mathbf{X})\,|$ of the number of neurons with $S_i=+1$ at time t in $\Omega(\mathbf{X})$:

$$p(t;\mathbf{X})\equiv\lim_{N\to\infty}\frac{|\,\{i\,|\,i\in\Omega(\mathbf{X}),S_i(t)=1\}\,|}{|\,\Omega(\mathbf{X})\,|}\,.$$

Then, we can readily obtain, in the thermodynamical limit, a one-body nonlinear master equation[3] for the time evolution of $p(t;\mathbf{X})$ from Eq. (2):

$$p(t+1;\mathbf{X})=\tfrac{1}{2}\{1+\tanh[\beta h(\mathbf{X},t)]\}\,, \qquad (4)$$

where

$$h(\mathbf{X},t)=\sum_{\mu\nu}A_{\mu\nu}\xi^{(\mu)}[m^{(\nu)}(t)+km^{(\nu)}(t-1)]$$
$$+\sum_{\mu\nu}A_{\mu\nu}[m^{(0)}(t)+km^{(0)}(t-1)]\,.$$

In the thermodynamical limit, the overlaps can be expressed in terms of the $p(t;\mathbf{X})$ as $m^{(\mu)}(t)=\sum_{\mathbf{X}}r(\mathbf{X})\times\xi^{(\mu)}[2p(t;\mathbf{X})-1]$ and $m^{(0)}(t)=\sum_{\mathbf{X}}r(\mathbf{X})\eta[2p(t;\mathbf{X})-1]$. Now, we can easily obtain a self-consistent nonlinear mapping describing the time evolution of the overlaps in a $(p+1)$-dimensional configuration space by using Eq. (4) describing the time evolution of 2^{p+1} probabilities $p(t;\mathbf{X})$:

$$m^{(\mu)}(t+1)=\sum_{\mathbf{X}\in\mathbf{H}^{p+1}}r(\mathbf{X})\xi^{(\mu)}\tanh\left[\beta\left(\sum_{a\gamma}A_{a\gamma}\xi^{(a)}[m^{(\gamma)}(t)+km^{(\gamma)}(t-1)]+\sum_{a\gamma}A_{a\gamma}[m^{(0)}(t)+km^{(0)}(t-1)]\right)\right],$$
$$\mu=1,\ldots,p\,, \qquad (5)$$
$$m^{(0)}(t+1)=\sum_{\mathbf{X}\in\mathbf{H}^{p+1}}r(\mathbf{X})\eta\tanh\left[\beta\left(\sum_{a\gamma}A_{a\gamma}\xi^{(a)}[m^{(\gamma)}(t)+km^{(\gamma)}(t-1)]+\sum_{a\gamma}A_{a\gamma}[m^{(0)}(t)+km^{(0)}(t-1)]\right)\right].$$

We see that specifying an excitatory or inhibitory character for each neuron in the present learning scheme results in the increase of the dimension of the mapping by 1. Owing to the reduction of the number of dynamical variables from 2^{p+1} to $p+1$, the analysis of the macroscopic dynamics (5) is much easier than that of Eq. (4).

In the present paper, we assume that $\{\xi_i^{(\mu)}\}$ and $\{\eta_i\}$ are independently specified: $r(\mathbf{X})=r(\xi)r(\eta)$, $r(\eta=1)\equiv r_e$. Then, the summations over η are taken in Eq. (5) to yield a trivial factor of 1 for the right-hand side of the first equation and $2r_e-1$ for that of the second.

We are interested in the behavior of the nonlinear mapping (5) in various parameter regions. We note that $m^{(0)}$ is 0 at an arbitrary instant when $r_e=\frac{1}{2}$ and is effectively decoupled from the other variables. Then, the

remaining mapping for $m^{(\mu)}$'s in Eq. (5) coincides with (discrete-time version of) that obtained in Ref. 3, in which the Dale hypothesis is not taken into account. Therefore the nontrivial effects of incorporating the Dale hypothesis can be seen only when the two types of neurons are not balanced in number in the case that $\{\xi_i^{(\mu)}\}$ and $\{\eta_i\}$ are independently specified.

In the high-temperature limit $\beta \to 0$, the right-hand sides of Eq. (5) are approximately linear in $m^{(\mu)}$'s and $m^{(0)}$ and the successive images of the mapping end up with a trivial fixed point $m=0$, which is a natural consequence of an infinitely large amount of thermal noise.[3] On the other hand, in the low-temperature limit $\beta \to \infty$, the hyperbolic tangent functions turn into the sign functions and the attractors of the mapping are at most limit cycles since the finiteness of p implies that the overlaps are allowed to take their values only in a set of finite elements. At the intermediate temperatures, the system is expected to show complicated behavior even when p is small.

We have conducted numerical simulations of Eq. (5) to explore the temperature dependence of the dynamical behavior of the pattern overlaps, assuming, for simplicity, that only two independent patterns are embedded [i.e., $p=2$, $r(\xi)=r(\xi^{(1)})r(\xi^{(2)})$]. We have found that there exist the parameter regions in which the system exhibits chaotic behavior.

We depict two examples of the chaotic attractors exhibited by the network systems with $k=0.8$ and $r_e<0.5$ (i.e., the network is dominated by inhibitory neurons). Figure 1 shows the projected phase portrait in the $m^{(1)}$-$m^{(2)}$ plane for the chaotic motion ($A_{11}=A_{12}=A_{22}=1$, $A_{21}=0$, $r_e=0.23$, $\beta=2.9$). Figure 2 displays another example of the chaotic retrieval ($A_{11}=A_{22}=1$, $A_{12}=4$, $A_{21}=0$, $r_e=0.24$, $\beta=2.95$) in which the retrieval trajectory becomes narrow in the $m^{(2)}$ direction since the retrieval motion towards $\{\xi_i^{(2)}\}$ always switches over that towards $\{\xi_i^{(1)}\}$, owing to the larger value of the off-diagonal element A_{12}. Chaotic behavior of nonlinear dynamical systems is studied by measuring the irregularity of the deterministic motions.[9] Positivity of the largest Lyapunov exponent, which measures the sensitivity to initial conditions, is the very evidence of the chaotic motions. The largest Lyapunov exponents have been calculated to be 0.07 and 0.26 for the chaotic attractors shown in Figs. 1 and 2, respectively, through 10^5 iterations of Eq. (5). The chaotic motions are also indicated by the appearance of broad noise in the power spectrum, an example of which is given in Fig. 3 for the chaotic attractor shown in Fig. 2.

We have found that as β is decreased from infinity, the system undergoes successively several types of bifurcations including the Ruelle-Takens-Newhouse route to chaos as well as the intermittency route,[9] until it exhibits relaxation motions towards a trivial fixed point $m=0$ at sufficiently small β. Since, as previously noted, the chaotic behavior in our neural-network systems manifests itself only at the intermediate temperature, i.e., only in the presence of finite magnitude of the external noise, we refer to such chaos as noise-driven chaos.

Without our numerical studies with β varied, we have observed that chaotic motions hardly appear when excitatory neurons dominate in the neural networks. Parisi suggested[10] that the inhibition may be responsible for driving the system's motions into chaotic ones in neural networks with random couplings. Our analysis, which is based on the properly learned synaptic couplings according to the Dale hypothesis and on a theoretically unambiguous method, supports this observation.

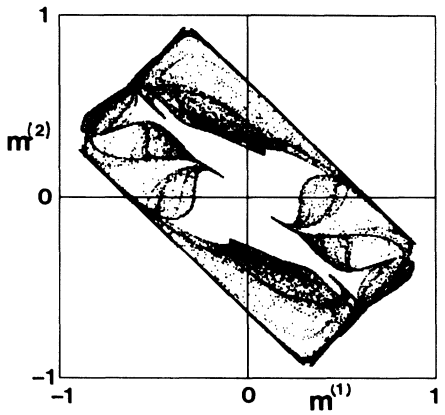

FIG. 1. Chaotic attractor at $\beta=2.9$ for $r_e=0.23$ is shown by the projected phase portrait in the $m^{(1)}$-$m^{(2)}$ plane [10^5 iterations of the map (5)]. The parameters are $r(\xi^{(1)}=1)=0.2$, $r(\xi^{(2)}=1)=0.8$, $A_{11}=A_{12}=A_{22}=1$, $A_{21}=0$, and $k=0.8$. The largest Lyapunov exponent has been numerically calculated to be 0.07. The chaotic attractor shown here has been found to coexist with a torus 1 to which a trajectory starting with different initial conditions settles in.

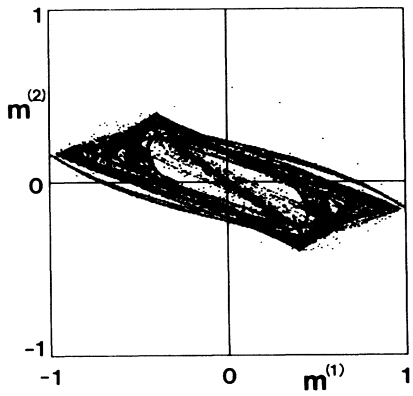

FIG. 2. Chaotic attractor at $\beta=2.95$ for $r_e=0.24$ (10^5 iterations) when the off-diagonal element A_{12} is comparatively large. The parameters are $r(\xi^{(1)}=1)=0.3$, $r(\xi^{(2)}=1)=0.7$, $A_{11}=A_{22}=1$, $A_{12}=4$, $A_{21}=0$, and $k=0.8$. The largest Lyapunov exponent is 0.26.

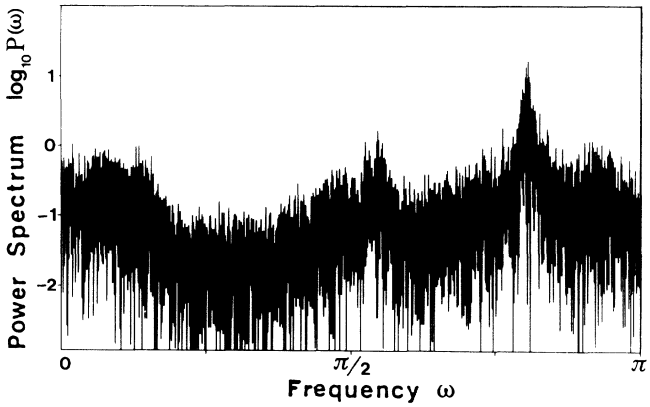

FIG. 3. The power spectrum of the chaotic motion shown in Fig. 2 which is obtained from the fast Fourier transform of the time series of 8192 points of the pattern overlap $m^{(1)}(t)$:

$$P(\omega)=\left|\frac{1}{\sqrt{T}}\sum_{t=0}^{T-1}m^{(1)}(t)\exp(-i\omega t)\right|^2,$$

with $\omega=2\pi n/T$ ($T=8192$, n is an integer). The continuous spectrum indicates chaotic behavior of the retrieval dynamics.

In summary, by undertaking a systematic approach based on the nonlinear master equation to the analysis of the dynamical behavior of asymmetric neural networks with the physiological Dale hypothesis taken into account, we have found that the noise-driven chaos can manifest itself under the simple non-Markovian dynamics of the Little type governing the networks with a finite number of embedded patterns. To our knowledge, the occurrence of chaos is observed for the first time *in associative neural networks with a finite number of embedded patterns.* The results may suggest the nontriviality of the Dale hypothesis as well as the effect of noise and encourage us to study further their inherent roles in order to understand the meanings and functions of chaos in the global activities of the physiological nervous systems.

[1]S. Amari, IEEE Trans. Comput. **21**, 1197 (1972).

[2]J. Buhmann and K. Schulten, Europhys. Lett. **4**, 1205 (1987); A. C. C. Coolen and Th. W. Ruijgrok, Phys. Rev. A **38**, 4253 (1988).

[3]M. Shiino, H. Nishimori, and M. Ono, J. Phys. Soc. Jpn. **58**, 763 (1989).

[4]J. C. Eccles, *Physiology of Synapses* (Springer-Verlag, Berlin, 1964).

[5]S. Shinomoto, Biol. Cybernet. **57**, 197 (1987).

[6]T. Fukai (to be published).

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

[8]M. Y. Choi and B. A. Huberman, Phys. Rev. B **28**, 254 (1983).

[9]J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).

[10]G. Parisi, J. Phys. A **19**, L675 (1986).

[11]I. Tsuda, E. Koerner, and H. Shimizu, Prog. Theor. Phys. **78**, 51 (1987); H. Sompolinsky, A. Crisanti, and H. J. Sommers, Phys. Rev. Lett. **61**, 259 (1988).

[12]H. Hayashi, S. Ishizuka, M. Ota, and K. Hirakawa, Phys. Lett. **88A**, 435 (1982); K. Aihara, G. Matsumoto, and M. Ichikawa, Phys. Lett. **111A**, 251 (1985).

[13]W. J. Freeman, Biol. Cybernet. **56**, 139 (1987); C. A. Skarda and W. J. Freeman, Behav. Brain Sci. **10**, 161 (1987).

[14]J. L. van Hemmen, Phys. Rev. Lett. **49**, 409 (1982); U. Riedel, R. Kühn, and J. L. van Hemmen, Phys. Rev. A **38**, 1105 (1988). In the latter, an analog of Eq. (4) is briefly noted.

[15]D. Kleinfeld, Proc. Natl. Acad. Sci. U.S.A. **83**, 9469 (1986); D. Kleinfeld and H. Sompolinsky, Biophys. J. **54**, 1039 (1988).