

## Statistical Mechanics of Temporal Association in Neural Networks with Transmission Delays

A. V. M. Herz,<sup>(1)</sup> Z. Li,<sup>(2)</sup> and J. L. van Hemmen<sup>(3)</sup>

<sup>(1)</sup>*Division of Chemistry, 139-74, California Institute of Technology, Pasadena, California 91125*

<sup>(2)</sup>*School of Natural Sciences, Institute for Advanced Study, Princeton, New Jersey 08540*

<sup>(3)</sup>*Physik-Department der Technische Universität München, D-8046 Garching, Germany*

(Received 4 September 1990)

We study the representation of static patterns and temporal sequences in neural networks with signal delays and a stochastic parallel dynamics. For a wide class of delay distributions, the asymptotic network behavior can be described by a generalized Gibbs distribution, generated by a novel Lyapunov functional for the deterministic dynamics. We extend techniques of equilibrium statistical mechanics so as to deal with time-dependent phenomena, derive analytic results for both retrieval quality and storage capacity, and compare them with numerical simulations.

PACS numbers: 87.10.+e, 64.60.Cn, 89.70.+c

Associative reconstruction of noisy data in a neural network can be accomplished by endowing a pattern to be memorized with a basin of attraction for the system's retrieval dynamics.<sup>1,2</sup> If the dynamics is governed by a Lyapunov function, a simple intuitive understanding of the global computation becomes possible: The system performs a downhill motion in an energy landscape created by the stored information.

For networks with static patterns, there is a Lyapunov function if the interactions between single neurons are instantaneous and mediated by symmetric couplings.<sup>2</sup> Methods of equilibrium statistical mechanics may then be applied and permit a quantitative analysis of the network's performance in terms of the retrieval quality and storage capacity.<sup>3</sup> The existence of a Lyapunov function is thus of great conceptual as well as technical importance.

In general, external inputs to a neural network are not limited to static memories but provide information in both space and time. To code the temporal aspects of sequences of patterns to be learned, additional asymmetric couplings with transmission delays may be introduced.<sup>4,5</sup> This approach can be generalized to networks with a broad distribution of signal delays where Hebb's neurophysiological principle for learning<sup>6</sup> naturally leads to a *joint* representation of spatial *and* temporal information.<sup>7</sup> However, no description in terms of a Lyapunov function has been given so far.

In this Letter such a description is developed for a certain class of networks with transmission delays. We present a Lyapunov functional for the deterministic parallel dynamics, generalize the formalism of equilibrium statistical mechanics so as to deal with thermal noise in systems with delayed interactions, and, hence, make the domain of time-dependent phenomena accessible to powerful free-energy techniques.

We follow Refs. 1-3 and model single neurons by Ising spins  $S_i$ ,  $1 \leq i \leq N$ . They represent a firing state for  $S_i = +1$  and a quiescent one for  $S_i = -1$ . The neurons

are connected by synapses with modifiable efficacies  $J_{ij}(\tau)$ . Here  $\tau$  denotes a fixed time delay for the information transport from  $j$  to  $i$ . We focus on a solitonlike propagation of neural signals, characteristic for the (axonal) transmission of action potentials, and consider a model where each pair of neurons is linked by several axons with delays  $0 \leq \tau \leq \tau_{\max}$ . External stimuli are fed into the system via two-state receptors  $\sigma_i = \pm 1$ . The local fields (postsynaptic potentials) are then given by

$$h_i(t) = (1 - \gamma) \sum_{j=1}^N \sum_{\tau=0}^{\tau_{\max}} J_{ij}(\tau) S_j(t - \tau) + \gamma \sigma_i(t), \quad (1)$$

where the parameter  $\gamma$  measures the system's input sensitivity. Other transport mechanisms and network architectures have been studied elsewhere;<sup>7</sup> see also Refs. 4, 5, 8, and 9.

We concentrate on synchronous (parallel) dynamics and set the basic time step equal to unity. Consequently, the signal delays take non-negative integer values. The influence of synaptic noise may be described by a stochastic Glauber dynamics<sup>10</sup> at finite inverse temperature  $\beta = T^{-1}$ . In that case,

$$\text{Prob}[S_i(t+1) = \pm 1] = \frac{1}{2} \{1 \pm \tanh[\beta h_i(t)]\}, \quad (2)$$

which reduces in the deterministic limit,  $\beta \rightarrow \infty$ , to  $S_i(t+1) = \text{sgn}[h_i(t)]$ . The retrieval dynamics can be studied in the bulk limit  $N \rightarrow \infty$  if the number and length of the stored sequences remain finite.<sup>7-9</sup>

In the present Letter, we surpass the limits of this kind of analysis and show that an approach in terms of statistical mechanics is feasible for networks whose couplings exhibit an "extended synaptic symmetry,"<sup>11</sup>

$$J_{ij}(\tau) = J_{ji}(D - 2 - \tau), \quad (3)$$

for some (fixed) integer  $D \geq 2$ . Links with  $\tau \geq D - 1$  are assumed to vanish, i.e.,  $\tau_{\max} \leq D - 2$ .

The condition (3) can be met by simple training algorithms for the synaptic efficacies. As an example, we

present a linear Hebbian rule. During a learning session, a coupling strength  $J_{ij}(\tau)$  is altered after each time step by an amount proportional to the product of  $S_j(t-\tau)$  and  $S_i(t+1)$ —due to the delay  $\tau$  in (1) and the discrete dynamics (2) it takes  $\tau+1$  units of time until neuron  $j$  influences the *state* of neuron  $i$  through that specific synapse. After  $P$  learning sessions, labeled by  $\mu$  and each of duration  $D_\mu$ , we thus obtain

$$J_{ij}(\tau) = \varepsilon(\tau) \frac{1}{N} \sum_{\mu=1}^P \sum_{t_\mu=1}^{D_\mu} S_i(t_\mu+1) S_j(t_\mu-\tau). \quad (4)$$

The *a priori* weights  $\varepsilon(\tau)$  are non-negative, normalized such that  $\sum_{\tau=0}^{\tau_{\max}} \varepsilon(\tau) = 1$ , and may be used to define various network architectures.<sup>7</sup> For a “clamped” learning scenario with  $\gamma=1$  in (1), the system evolves strictly according to the external inputs,  $S_i(t_\mu) = \sigma_i(t_\mu-1)$ . Equation (3) holds once we teach *cyclic* associations with a common period  $D$  to networks whose weight distributions obey  $\varepsilon(\tau) = \varepsilon(D-2-\tau)$ .

Returning to the general case, given by (3), we now establish that there is a Lyapunov functional for the noiseless retrieval dynamics. We focus on a free evolution of the system, i.e., we set  $\gamma=0$  in (1), generalize the results of Ref. 12, and take

$$H(t) \equiv -\frac{1}{2} \sum_{i,j=1}^N \sum_{a=0}^{D-1} \sum_{\tau=0}^{\tau_{\max}} J_{ij}(\tau) S_i(t-a) \times S_j(t-(a+\tau+1)(\text{mod}D)). \quad (5)$$

The functional  $H$  depends on all network states between  $t+1-D$  and  $t$  so that solutions with constant  $H$  need not be static fixed points of the retrieval dynamics. The difference  $\Delta H(t) \equiv H(t) - H(t-1)$  is given by

$$\Delta H(t) = -\sum_{i=1}^N [S_i(t) - S_i(t-D)] h_i(t-1). \quad (6)$$

Deterministic dynamics, i.e.,  $S_i(t) = \text{sgn}[h_i(t-1)]$ , implies that the right-hand side will only take nonpositive values. For finite  $N$ ,  $H$  is bounded from below and  $\Delta H$  has to vanish for  $t \rightarrow \infty$ . This can be realized only if the system settles into a state with  $S_i(t) = S_i(t-D)$  for all  $i$ .

We have thus exposed two important facts: (a) The retrieval dynamics of a certain class of networks with delayed interactions is governed by a Lyapunov functional, and (b) the system relaxes to static states or limit cycles with  $S_i(t) = S_i(t-D)$ —oscillatory solutions with the same period as that of the taught sequences (or a period equal to an integer fraction of  $D$ ). It remains to be shown that such a limit cycle of the retrieval dynamics indeed resembles a stored sequence for appropriate initial conditions.

In order to prove the above assertion, we proceed in three steps. First, we demonstrate that our task concerning temporal associations with a cyclic structure can be mapped onto a symmetric network without signal delays.

Second, we apply equilibrium statistical mechanics to study such “*equivalent systems*.” Third, we focus on a special case, the Hebbian learning rule (4), and derive analytic results for the retrieval quality and storage capacity.

$D$ -periodic oscillatory solutions of the retrieval dynamics can be interpreted as static states in a “ $D$ -plicated” system. Let us therefore consider such a fictitious system with  $D$  columns and  $N$  rows of  $\pm 1$  neurons whose activities will be denoted by  $S_{ia}$ . A network state will be written  $A = (A_0, A_1, \dots, A_{D-1})$  with  $A_a \equiv \{S_{ia}; 1 \leq i \leq N\}$ . To reproduce the parallel dynamics of the original system, neurons  $S_{ia}$  with  $a = t(\text{mod}D)$  are updated at time  $t$ . The time evolution of the new network therefore has a pseudosequential characteristic: synchronous within single columns and sequentially ordered with respect to these columns. Accordingly, the neural activities at time  $t$  are given by  $S_{ia}(t) \equiv S_i(a+n_t)$  for  $a \leq t(\text{mod}D)$  and  $S_{ia}(t) \equiv S_i(a+n_t-D)$  for  $a > t(\text{mod}D)$ , where  $n_t$  is defined through  $t \equiv n_t + t(\text{mod}D)$ . Because of (3), we may construct *symmetric* efficacies  $J_{ij}^{ab} = J_{ji}^{ba}$  for the new system by putting  $J_{ij}^{aa} = 0$  and

$$J_{ij}^{ab} = J_{ij}((b-a-1)(\text{mod}D)) \quad \text{for } a \neq b, \quad (7)$$

so that there is a well-defined Hamiltonian,

$$H = -\frac{1}{2} \sum_{i,j=1}^N \sum_{a,b=0}^{D-1} J_{ij}^{ab} S_{ia} S_{jb}, \quad (8)$$

equal to that of a Hopfield network of size  $ND$ .

An evaluation of (8) in terms of the former state variables reveals that it is identical with the Lyapunov functional (5). We may once again calculate  $\Delta H(t)$  and arrive at a formula corresponding to (6). The interpretation, however, is changed: a *limit cycle* of period  $D$  in the original network corresponds to a *fixed point* of the new system of size  $ND$ . We have thus shown that the deterministic parallel dynamics of a delay network with extended symmetry can be understood in terms of a downhill motion in the energy landscape of its equivalent system.

If, in a second step, we now switch to a stochastic Glauber dynamics (2), the important question arises whether  $H$  also determines the invariant distribution  $\rho$  of the network. This need not be true since the columnwise updating procedure of the equivalent system differs from both the Little<sup>1</sup> and the Hopfield<sup>2</sup> model. It turns out, however, that there is indeed an equilibrium distribution *à la* Gibbs,<sup>11</sup>

$$\rho(A) = Z^{-1} \exp[-\beta H(A)]. \quad (9)$$

Here  $Z \equiv \text{Tr}_A \exp[-\beta H(A)]$  denotes the partition function. The exact proof is elaborate since due to the non-Markovian structure of the dynamics (1) and (2), the principle of detailed balance does not apply to single-step transition probabilities.

In passing we note that for  $D=2$  there are only links

with zero delay. By (3) we have  $J_{ij}(0) = J_{ji}(0)$ , i.e., we are dealing with a symmetric Little model. We may introduce a reduced probability distribution  $\tilde{\rho}$  for this special case,  $\tilde{\rho}(A_1) \equiv \text{Tr}_{A_0 \rho}(A_0 A_1)$ , and obtain  $\tilde{\rho}(A_1) = Z^{-1} \exp[-\beta \tilde{H}(A_1)]$  with

$$\tilde{H} \equiv -\beta^{-1} \sum_{i=1}^N \ln \left[ 2 \cosh \left[ \beta \sum_{j=1}^N J_{ij} S_j \right] \right]. \quad (10)$$

We thus have recovered both the effective Hamiltonian of the Little model as derived by Peretto<sup>10</sup> and the duplicated-system ( $D=2$ ) technique of Ref. 13.

Finally, we finish our argument by turning to quantitative results. We focus on synaptic couplings generated by the Hebbian rule (4) and investigate the case where each of the  $P$  learning sessions corresponds to teaching a (different) cycle of  $D$  patterns  $\{\psi_{ip}^{\mu}; 1 \leq i \leq N\}$ ,  $1 \leq p \leq D$ , i.e., each pattern lasts for a single time step,  $\sigma_i(t_{\mu}) = \psi_{i,t_{\mu}(\text{mod } D)}$ . In accordance with (4) and (8), there are  $PD$  patterns  $\{\xi_{ia}^{\mu}; 1 \leq i \leq N, 0 \leq a \leq D-1\}$  for the equivalent network,  $1 \leq \mu \leq P$ ,  $1 \leq a \leq D$ , given by  $\xi_{ia}^{\mu 0} = \psi_{ia}^{\mu}$  and  $\xi_{ia}^{\mu \alpha} = \xi_{i,(a-\alpha)(\text{mod } D)}^{\mu 0}$ . In what follows, we work with unbiased random patterns where  $\psi_{ip}^{\mu} = \pm 1$  with equal probability, and study our network at a finite storage level  $\alpha = \lim_{N \rightarrow \infty} (P/N) > 0$ .

As in the replica-symmetric theory of Amit, Gutfreund, and Sompolinsky<sup>3</sup> we single out finitely many,

$$-\beta f(\beta) = -\frac{\beta}{2} \sum_{\nu} (m^{\nu})^2 - \frac{1}{2} \alpha \sum_k \left[ \ln [1 - \beta(1-q)\lambda_k(\mathcal{E})] - \frac{q\beta\lambda_k(\mathcal{E})}{1 - \beta(1-q)\lambda_k(\mathcal{E})} \right] + \frac{1}{2} \alpha \beta^2 (q-1)r + \left\langle \left\langle \ln 2 \cosh \left[ \beta \left\{ \sum_{\nu} m^{\nu} \xi^{\nu 0} + \sqrt{ar} z \right\} \right] \right\rangle \right\rangle. \quad (16)$$

The preceding equations closely resemble their counterparts for the Hopfield (or Little) model and become identical to them for a "maximally uniform" distribution,  $\mathcal{E}_{ab} = (D-1)^{-1}(1 - \delta_{ab})$ , in the limit  $D \rightarrow \infty$ . For general period  $D$  and weight distribution  $\mathcal{E}$ , there is at zero temperature a first-order phase transition between the retrieval state and a spin-glass phase as in the Hopfield case. The critical storage level  $\alpha_c$  and the corresponding overlap  $m_c$  are different, though.

It should be noted that each cycle consists of  $D$  independent patterns so that the storage capacity for *single* patterns is  $\tilde{\alpha}_c = D\alpha_c$ . During the recognition process, however, each of them will trigger the cycle it belongs to and cannot be retrieved as a static pattern.

As a second measure of the performance, we include the information content  $I_R(\mathcal{E})$ , measured *per synapse* and *relative* to that of the Hopfield model,

$$I_R(\mathcal{E}) \equiv \frac{I(\mathcal{E})}{I(\text{Hopfield})} = \frac{D\alpha_c(\mathcal{E})}{d\alpha_c(\text{Hopfield})},$$

where  $d$  denotes the number of different delay lines between each pair of neurons,  $1 \leq d \leq D-1$ . A numerical

say  $s$ , cycles of the original system (corresponding to  $sD$  static patterns for the equivalent model of size  $ND$ ) and assume that the network is in a state highly correlated with these spatiotemporal objects. The remaining, extensively many cycles are described as a noise term. We consider *retrieval solutions* and make the ansatz

$$m_{\sigma a}^{\mu} \equiv N^{-1} \sum_i \xi_{ia}^{\mu} S_{ia}^{\sigma} = m^{\mu} \delta_{a,0}, \quad (11)$$

$$q_{ab}^{\rho\sigma} \equiv N^{-1} \sum_i S_{ia}^{\rho} S_{ib}^{\sigma} = \delta_{ab} [\delta_{\rho\sigma}(1-q) + q], \quad (12)$$

for the relevant order parameters of the  $n$ -fold replicated network,  $1 \leq \rho, \sigma \leq n$ . By applying the techniques of Refs. 3 and 14, we arrive at the fixed-point equations

$$m^{\mu} = \left\langle \left\langle \xi^{\mu 0} \tanh \left[ \beta \left\{ \sum_{\nu} m^{\nu} \xi^{\nu 0} + \sqrt{ar} z \right\} \right] \right\rangle \right\rangle, \quad (13)$$

$$q = \left\langle \left\langle \tanh^2 \left[ \beta \left\{ \sum_{\nu} m^{\nu} \xi^{\nu 0} + \sqrt{ar} z \right\} \right] \right\rangle \right\rangle, \quad (14)$$

$$r \equiv q \sum_{k=1}^D \frac{[\lambda_k(\mathcal{E})]^2}{[1 - \beta(1-q)\lambda_k(\mathcal{E})]^2}. \quad (15)$$

The double angular brackets represent an average with respect to both the condensed cycles and the normalized Gaussian random variable  $z$ . The  $\lambda_k(\mathcal{E})$  are eigenvalues of the matrix  $\mathcal{E}$  with  $\mathcal{E}_{ab} \equiv \varepsilon((b-a-1)(\text{mod } D))$ . At the saddle points given by (13)–(15), the free energy is

solution of the saddle-point equations at  $T=0$  leads to the following results for networks with a "maximally uniform" distribution:

$D$	$\alpha_c$	$m_c$	$I_R$
2	0.100	0.93	1.45
3	0.110	0.95	1.20
4	0.116	0.96	1.12
5	0.120	0.96	1.09

The above findings agree well with estimates from a finite-size analysis ( $N \leq 3000$ ) of data from numerical simulations as shown by two examples. For  $D=3$ , we have found  $\alpha_c = 0.120 \pm 0.015$ , and for  $D=4$ ,  $\alpha_c = 0.125 \pm 0.015$ .

Our results demonstrate that the storage capacity for temporal associations is comparable to that for static memories. As an example, we take  $D=2$ , i.e., the Little model. In the limit of large  $N$ , we see that  $0.100N$  two-cycles of the form  $\psi_{i1}^{\mu} \neq \psi_{i2}^{\mu}$  may be recalled as compared to  $0.138N$  static patterns;<sup>15</sup> this leads to a 1.45-fold in-

crease of the information content per synapse. At the same time, the retrieval overlap drops slightly from 0.97 to 0.93. The influence of the weight distribution on the network behavior may be demonstrated by some choices of  $\varepsilon(\tau)$  for  $D=4$ :

$\tau$	=	0	1	2	3	$\alpha_c$	$m_c$	$I_R$
$\varepsilon(\tau)$	=	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0	0.116	0.96	1.12
$\varepsilon(\tau)$	=	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0.100	0.93	1.45
$\varepsilon(\tau)$	=	0	1	0	0	0.050	0.93	1.45

The uniform distribution leads to the largest  $\alpha_c$  but smallest  $I_R$ . The other two networks have the same value of  $I_R$  as the (unique)  $D=2$  system due to the particular structure of their eigenvalue spectrum. Furthermore, one obtains  $I_R=1.45$  independently of  $D$  for all networks with a "minimal connectivity" where only one synapse links two neurons ( $D$  even). Simulation data show once again slightly higher values of  $\alpha_c$ —possibly indicating effects of replica-symmetry breaking as in the Hopfield model.<sup>3</sup>

In summary, we have introduced a Lyapunov functional for a wide class of neural networks with signal delays and have generalized equilibrium statistical mechanics so as to analyze the network behavior under a stochastic dynamics. Quantitative results have been obtained for a Hebbian learning rule. They prove that an extensive number of temporal associations can be stored as spatiotemporal attractors for the retrieval dynamics. Even at the critical storage level, retrieved sequences contain less than 3.5% errors for all studied architectures. Our results indicate that dynamical systems with delayed interactions can be programmed in a very efficient manner to perform associative computations in the space-time domain.

The current approach provides a new framework for the study of temporal associations in delayed networks. Though limited to networks whose couplings exhibit an "extended symmetry," it may be generalized in various directions. To mention a few examples, first, more sophisticated learning prescriptions like a spatiotemporal version of the pseudoinverse rule can be investigated. Second, one may wish to switch to continuous state variables and study temporal associations in "iterated-map networks."<sup>16</sup> It turns out that a generalization of the Lyapunov functional (5) covers that case as well,<sup>17</sup> al-

lowing for a direct comparison of theoretical predictions with experimental results from hardware implementations. Finally, one could try to develop a Lyapunov functional for a continuous-time dynamics including signal delays which seems to be rather significant for applications as well as for the general theory of functional differential equations and dynamical systems.

We thank J. J. Hopfield and R. Kühn for many stimulating discussions. Both of them as well as J. Buhmann and D. MacKay helped us with constructive comments on the manuscript. A.V.M.H. acknowledges support from the Studienstiftung des Deutschen Volkes. Z.L. is supported in part through a grant from the Seaver Institute.

<sup>1</sup>W. A. Little, *Math. Biosci.* **19**, 101 (1974).

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

<sup>3</sup>D. J. Amit, H. Gutfreund, and H. Sompolinsky, *Phys. Rev. A* **32**, 1007 (1985); *Ann. Phys. (N.Y.)* **173**, 30 (1987).

<sup>4</sup>D. Kleinfeld, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9469 (1986).

<sup>5</sup>H. Sompolinsky and I. Kanter, *Phys. Rev. Lett.* **57**, 2861 (1986).

<sup>6</sup>O. Hebb, *The Organization of Behavior* (Wiley, New York, 1949), p. 62.

<sup>7</sup>A. V. M. Herz, B. Sulzer, R. Kühn, and J. L. van Hemmen, *Europhys. Lett.* **7**, 663 (1988); *Biol. Cybern.* **60**, 457 (1989).

<sup>8</sup>A. C. C. Coolen and C. C. A. M. Gielen, *Europhys. Lett.* **7**, 281 (1988).

<sup>9</sup>M. Kerszberg and A. Zippelius, *Phys. Scr.* **T33**, 54 (1990).

<sup>10</sup>P. Peretto, *Biol. Cybern.* **50**, 51 (1984).

<sup>11</sup>A. V. M. Herz, Ph.D. thesis, Heidelberg, 1990 (unpublished); Z. Li and A. V. M. Herz, in *Neural Networks*, Proceedings of the Eleventh Sitges Conference, Lecture Notes in Physics Vol. 368, edited by L. Garrido (Springer-Verlag, Berlin, 1990), pp. 287–302.

<sup>12</sup>Z. Li (unpublished).

<sup>13</sup>J. L. van Hemmen, *Phys. Rev. A* **34**, 3435 (1986).

<sup>14</sup>J. L. van Hemmen and R. Kühn, *Phys. Rev. Lett.* **57**, 913 (1986); J. L. van Hemmen and V. A. Zagrebnov, *J. Phys. A* **20**, 3989 (1987).

<sup>15</sup>J. F. Fontanari and R. Köberle, *Phys. Rev. A* **36**, 2475 (1987).

<sup>16</sup>C. M. Marcus and R. M. Westervelt, *Phys. Rev. A* **42**, 2410 (1990).

<sup>17</sup>A. V. M. Herz (to be published).