

## Noise-Driven Temporal Association in Neural Networks.

J. BUHMANN and K. SCHULTEN

*Physik-Department, Technische Universität München, James-Frank-Str.  
8046 Garching, BRD.*

(received 24 June 1987; accepted in final form 26 August 1987)

PACS. 87.30. – Biophysics of neurophysiological processes.

**Abstract.** – A network of spinlike neurons with asymmetric exchange interactions and stochastic spike response which can learn and recall time sequences of biased patterns is proposed. Noise makes synapses with delayed response or with time-dependent strength, previously proposed for storage of time sequences, superfluous. An accurate timing of pattern sequences requires a sufficient number  $N$  of neurons. The performance of the suggested network is described by Monte Carlo simulation, in terms of a Fokker-Planck equation and, for  $N \rightarrow \infty$ , in terms of a Liouville equation.

Neural networks with spinlike neuronal units and symmetric synapses possess a highly structured phase space with many local minima in a global energy function. Little [1], Hopfield [2] and others [3, 4] have shown that such networks can be used as content-addressable memories. For this purpose the synaptic connections (exchange interactions  $J_{ik}$ ) are chosen such that the equilibrium states of the network coincide with states which represent stored *static* patterns. In this letter we address the problem of a content-addressable memory for *temporal* patterns. Previous attempts to construct such memories have relied on nonsymmetric synaptic interactions with temporal features: the synapses transmitted the interaction with a time delay [5-8] or the synapses changed their interactive strength in time [9, 10]. Such synapses have been observed only in very few and very specialized neural systems and, therefore, should not be invoked for temporal storage in brains.

We show below that spinlike networks can realize storage of temporal patterns with synaptic interactions which are neither time dependent nor exhibit a delay. The networks proposed in this letter require only two new features: 1) synaptic interactions must be sufficiently nonsymmetric, 2) noise must be present. Both features are considered to be natural attributes of biological networks. The asymmetric synapses in our model as in models suggested previously [5-9] provide directed projections between equilibrium states of the network and, thereby, define a sequential order among stored patterns. Noise triggers transitions between consecutive patterns. Because of this, stochastic attribute transition times are not precise. However, if stored patterns are represented by large enough sets of neurons, fluctuations in the transition time decrease.

Our network is composed of  $N$  neurons described by dynamic variables  $\{S_i\}_{i=1}^N$ . Neuron  $i$  is either firing ( $S_i = 1$ ) or quiet ( $S_i = 0$ ). The variables are updated asynchronously according to a probabilistic rule which represents the action of noise in the system:

$$S_i(t+1) = \begin{cases} 1, & \text{with probability } f_i(t), \\ 0, & \text{with probability } 1 - f_i(t), \end{cases} \quad (1)$$

where  $f_i = (1 + \exp[-(h_i - U)/T])^{-1}$  (<sup>1</sup>) for a molecular field  $h_i = \sum_k J_{ik} S_k$ . The parameters  $U$  and  $T$  are the threshold potential and the network temperature.

The patterns  $S^\nu = \{S_i\}_{i=1}^N$  which we intend to store in our network are biased, *i.e.* consist of only a small fraction of firing neurons with  $S_i^\nu = 1$  without the requirement of statistical independence of the  $S_i^\nu$ . This allows to store meaningful, *i.e.* nonrandom patterns. Storage of patterns with low level of activity (see also [11]) appears to be closer to the behaviour found in brains. We will assume the orthogonality condition  $\sum_i \varepsilon^\nu S_i^\nu S_i^\mu = \delta_{\nu\mu}$ .

Storage is achieved in our network by a choice of exchange interactions which differ from that of the Hopfield network. The exchange interaction  $J_{ik}$  is composed of a symmetric and an antisymmetric part. The symmetric interaction between neuron  $k$  and  $i$

$$J_{ik}^0 = \sum_{\nu=1}^m \left( \varepsilon^\nu S_k^\nu S_i^\nu - \sum_{\substack{\mu=1 \\ \mu \neq \nu}}^m \gamma \frac{m}{N} S_k^\mu S_i^\nu \right)$$

consists of an excitatory and an inhibitory contribution. The first (excitatory) term induces a cooperation between neurons which represent the same pattern  $\nu$ . The second (inhibitory) term gives rise to a competition between all the stored patterns ( $0.1 < \gamma < 1.0$ ). The resulting dynamics selects asymptotically a network state which coincides with that pattern with the largest overlap with the initial state.

In order to retrieve the patterns  $S^\nu$  in the sequence  $\nu = 1, 2, \dots, m$ , we modify the symmetric exchange interaction  $J_{ik}^0$ . We first eliminate the weak inhibition between predecessor-successor patterns in the sequence, *i.e.* between  $S^\nu$  and  $S^{\nu+1}$ , and add an excitatory (positive) projection from  $S^\nu$  to  $S^{\nu+1}$  and an inhibitory backward-projection from  $S^{\nu+1}$  to  $S^\nu$ :

$$J_{ik} = \sum_{\nu=1}^m \left( \varepsilon^\nu S_k^\nu - \sum_{\substack{\mu=1 \\ |\nu-\mu|>1}}^m \gamma \frac{m}{N} S_k^\mu + \alpha^{\nu-1} \varepsilon^{\nu-1} S_k^{\nu-1} - \beta^{\nu+1} \varepsilon^{\nu+1} S_k^{\nu+1} \right) S_i^\nu. \quad (2)$$

If the network is in state  $S^\nu$ , the obvious effect of the added excitatory and inhibitory contributions is to excite  $S^{\nu+1}$ , and when this state is sufficiently presented to inhibit  $S^\nu$ . This induces the consecutive retrieval of patterns  $S^\sigma, S^{\sigma+1}, \dots, S^m$  when the network starts in a state exhibiting dominant overlap with pattern  $S^\sigma$ . We found parameter values  $U = 0.35$  and  $T = 0.1$  suitable to realize such retrieval.

A salient feature in the choice of the exchange interaction (2) and the orthogonality condition  $\sum_i \varepsilon^\nu S_i^\nu S_i^\mu = \delta_{\nu\mu}$  is that the molecular field depends solely on the *magnetizations*

(<sup>1</sup>) The theory presented below requires only a sigmoidal shape of the function  $f_i$ , *i.e.* it does not depend on the analytical expression given here.

$x^v(t) = \sum_k \epsilon^v S_k^v S_k(t)$  which measure the overlap of pattern  $S^v$  with the momentaneous network state  $S(t)$ . The network, therefore, can be completely described by the dynamic variables  $x^v(t), v = 1, \dots, m$  and the probability  $f_i$  in (1) can be replaced by

$$f^v(x(t)) = \left( 1 + \exp \left[ -\frac{\mathcal{E}^v - U}{T} \right] \right)^{-1}, \tag{3}$$

$$\mathcal{E}^v = \alpha^{v-1} x^{v-1} + x^v - \beta^{v+1} x^{v+1} - \sum_{\substack{\mu=1 \\ |\nu-\mu|>1}}^m \gamma \frac{m}{N} \epsilon^v x^\mu,$$

where  $\mathcal{E}^v$  is the molecular field of all neurons  $i$  for which  $S_i^v = 1$  holds. (Neurons which are not active in any stored pattern according to (2) are strongly inhibited by a molecular field  $\mathcal{E} = -\sum \gamma m / \epsilon^\mu N x^\mu$ .) The introduction of the variables  $x^v(t)$  drastically reduces the dimension of the phase space of the network, namely from  $2^N$  to  $m$ . This reduction results from the fact that the parameters  $\alpha, \beta, \gamma, U, T$  in (1), (3) do not depend on the neuron index  $i$ . In case of networks with more heterogeneous interaction parameters, our model can also be applied and corresponds then to a mean-field approximation.

The result of a Monte Carlo simulation of our model is shown in fig. 1. The network considered has stored 8 patterns and is described by  $x^v(t), v = 1, \dots, 8$ . The functions  $x^v(t)$  each are seen to assume small resting values except for a brief period when the  $x^v$  rise close to the value one. The ordering of these periods implies that the network jumps consecutively from one pattern to the next. As an illustration we present in fig. 2 a second Monte Carlo simulation of a network which can retrieve a sequence of six patterns representing the numbers 1, 2, ..., 6. The second and fourth network states show the network during the brief periods when patterns are switched.

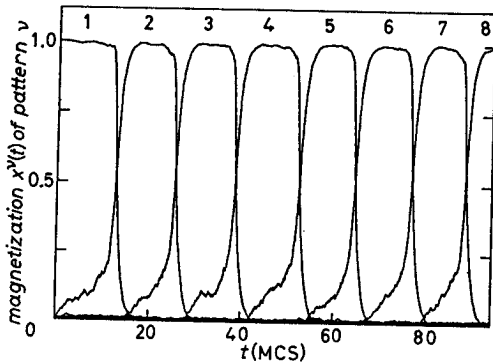


Fig. 1.

Fig. 1. - Simulation of the network with the parameters  $v=8, \epsilon^v=0.001 \forall v, \alpha^v=0.1, \beta^v=1.0, T=0.1, U=0.35$ . The curves show the evolution of the magnetizations  $x^v(t)$  as a function of time.

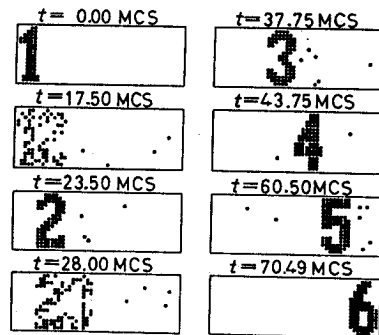


Fig. 2.

Fig. 2. - Network activity monitored over a Monte Carlo simulation with parameters  $\alpha^v=0.1, \beta^v=1.0 \forall v, T=0.1, U=0.35$ . The network recalls the stored pattern sequence 1, 2, ..., 6.

To study the dynamics of our network, we derive a master equation which describes the probability that the network assumes specific  $x^v$  values at time  $t$ . A closed and simple master equation in terms of  $x^v$  can be derived since the evolution of the network depends only on the magnetizations  $\mathbf{x} = \{x^v(t)\}_{v=1}^m$  and since the asynchronous updates (1) affect always only a single  $x^v(t)$ .

Because of the asynchronous update of neurons, different magnetizations  $x^\nu$  are coupled only through the spike probability  $f^\nu$  given in eq. (3). (Note that  $x^\nu$  can assume the values  $0, \epsilon^\nu, 2\epsilon^\nu, \dots, 1$ .) Summing up processes which increase or decrease the probability  $p(\mathbf{x}, t)$  to find the system in a state with magnetization  $x^\nu$ , one obtains the master equation

$$\frac{\tau}{N} \frac{\partial}{\partial t} p(\mathbf{x}, t) = \sum_{\nu=1}^m \frac{1}{\epsilon^\nu N} [p(x^\nu + \epsilon^\nu)(x^\nu + \epsilon^\nu)(1 - f^\nu(x^\nu + \epsilon^\nu)) + p(x^\nu - \epsilon^\nu)(1 - x^\nu + \epsilon^\nu)f^\nu(x^\nu - \epsilon^\nu) - p(x^\nu)(x^\nu + f^\nu(x^\nu) - 2x^\nu f^\nu(x^\nu))], \quad (4)$$

where the arguments  $x^\mu, \mu \neq \nu$  of  $p(\mathbf{x}, t)$  and of  $f^\nu(\mathbf{x}, t)$  which remain unaltered are suppressed on the r.h.s.. The first term on the r.h.s describes the transition of a firing neuron belonging to pattern  $\nu$  to the resting state ( $S_i(t) = 0$ ), the second term describes the inverse process. The last term accounts for processes which decrease the probability to find the magnetization  $x^\nu$ . The asynchronous updates of neurons establish a characteristic time scale of  $\tau/N$  for one spin flip, where  $\tau$  corresponds to one Monte Carlo step (MCS).

In the limit of large patterns ( $\epsilon^\nu \rightarrow 0$ ) the discrete master equation becomes a continuous equation in  $\mathbf{x}$ . Taylor expansion of (4) to order  $(\epsilon^\nu)^2$  results in the Fokker-Planck equation

$$\tau \frac{\partial}{\partial t} p(\mathbf{x}, t) = \sum_{\nu=1}^m \left[ -\frac{\partial}{\partial x^\nu} F^\nu p + \frac{\epsilon^\nu}{2} \frac{\partial^2}{\partial x^{\nu 2}} D^\nu p \right], \quad (5)$$

with  $F^\nu = -(x^\nu - f^\nu)$  and  $D^\nu = x^\nu + f^\nu - 2x^\nu f^\nu$ . In the limit of infinitely large patterns ( $\epsilon^\nu = 0$ ) eq. (5) simplifies further and assumes the form of a deterministic Liouville equation [12]. In this limit the vector of magnetizations  $\mathbf{x}$  obeys the kinetic equation [13]

$$\tau \frac{d}{dt} x^\nu = -(x^\nu - f^\nu(\mathbf{x})). \quad (6)$$

The threshold  $U$  in (3) has to prevent fluctuations from activating patterns in wrong succession, but should not be as large as to suppress all pattern states during a transition from pattern  $\nu$  to  $\nu + 1$ , when  $x^\nu$  and  $x^{\nu+1}$  assume values around 0.5. To satisfy these requirements,  $U$  has to be in the range  $T < U < 0.5$ . Inserting such  $U$  into (3) yields  $f^\nu \approx 1$  and  $f^\mu \approx 0, \mu \notin \{\nu, \nu + 1\}$  when the network occupies the  $\nu$ -th pattern state, *i.e.* for  $x^\mu = \delta_{\mu\nu}$ . These considerations allow us to investigate the transition from pattern  $\nu$  to  $\nu + 1$  by projecting the dynamics onto the plane spanned by  $\{x^\nu, x^{\nu+1}\}$ . In fig. 3 the magnetizations  $x^\nu(t)$  and  $x^{\nu+1}(t)$  corresponding to the solution of (6) for the initial state  $(x^\nu(0), x^{\nu+1}(0)) = (1, 0)$  are compared with the magnetizations obtained from a Monte Carlo simulation.

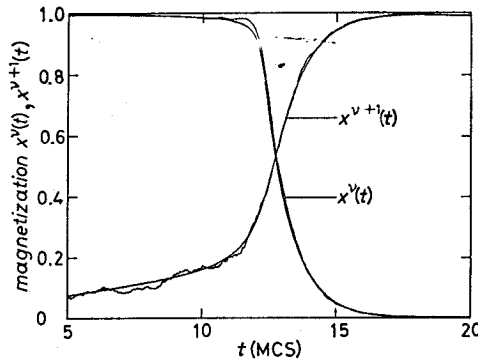


Fig. 3. - Time dependence of magnetizations  $x^\nu(t)$  and  $x^{\nu+1}(t)$ . The solution of (6) is compared with magnetizations received from Monte Carlo simulations.

The transition time  $\tau_{\nu \rightarrow \nu+1}$  necessary for the network to evolve from pattern state  $\nu$  to pattern state  $\nu + 1$  can be calculated analytically [14]. The derivation will be published elsewhere. The transition time increases with decreasing projection strength  $\alpha'$  and diverges for a critical temperature  $T^*$  defined by the condition that the minimal force  $F^{\nu+1}|_{x^{\nu}=1}$  which drives the network towards the succeeding pattern  $\nu + 1$  vanishes. For temperatures below  $T^*$ , the pattern state  $\nu$  remains stable for an infinite time. In the limit  $\varepsilon^{\nu} \rightarrow 0$  one can actually tune the network to stay for an arbitrary time in the pattern state  $\nu$  and then to jump to the next pattern  $\nu + 1$ .

The analysis above for infinite networks can serve as an approximation for finite networks [12]. For this purpose (5) can be transformed to variables  $y^{\nu}/\varepsilon^{\nu}$ , where the  $y^{\nu}$  measure the deviation from the solution of (6). With increasing pattern size the time variance  $\sigma^{\nu}$  of a single transition decreases as  $\sqrt{\varepsilon^{\nu}}$ . Another consideration concerns the question in how far patterns may overlap and still be properly recalled. An overlap is tolerable as long as it conserves the recall order of the patterns. However, such overlap can strongly affect the transition time between consecutive patterns. The dynamics of a network with nonorthogonal patterns stored can be described in terms of the enlarged set of variables  $x^{\nu_1, \dots, \nu_k}(t)$  which count all neurons active at time  $t$  and belonging to all of the patterns  $\nu_1, \dots, \nu_k$ . An analysis of the resulting dynamics shows that an overlap between consecutive patterns corresponds to an enhanced projection strength from  $\nu$  to  $\nu + 1$ . An appropriate choice [14] of  $\alpha'$  allows to tune the transition time  $\tau_{\nu \rightarrow \nu+1}$  for a given overlap  $x^{\nu, \nu+1}$ .

We have demonstrated that networks of spinlike neurons can store and recall time sequences of patterns by means of nonsymmetric, time-independent and instantaneous exchange interactions. The transition between patterns during recall is triggered by global noise. The fluctuations of the transition times decrease with increasing network size. Without demonstration we like to point out that the performance of the storage device could be enhanced if the global noise is made time dependent [14].

\* \* \*

We wish to thank R. DIVKO and H. RITTER for helpful discussions.

## REFERENCES

- [1] LITTLE W. A., *Math. Biosci.*, 19 (1974) 101; LITTLE W. A. and SHAW G. L., *Behav. Biol.*, 14 (1975) 115.
- [2] HOPFIELD J. J., *Proc. Natl. Acad. Sci. U.S.A.*, 79 (1982) 2554; 81 (1984) 3088.
- [3] AMIT D. J., GUTFREUND H. and SOMPOLINSKY H., *Phys. Rev. A*, 32 (1985) 1007; *Phys. Rev. Lett.*, 55 (1985) 1530.
- [4] KINZEL W., *Z. Phys. B*, 60 (1985) 205.
- [5] SOMPOLINSKY H. and KANTER I., *Phys. Rev. Lett.*, 57 (1986) 2861.
- [6] KLEINFELD D., *Proc. Natl. Acad. Sci. U.S.A.*, 83 (1986) 9469.
- [7] KLEINFELD D. and SOMPOLINSKY H., preprint (1987).
- [8] BUHMANN J., Diplom-Thesis, Technische Universität München, 1985.
- [9] PERETTO P. and NIEZ J. J., in *Disordered Systems and Biological Organization*, edited by E. BIENENSTOCK, (Springer Verlag, Berlin), 1986.
- [10] DEHAENE S., CHANGEUX J.-P. and NADAL J.-P., *Proc. Natl. Acad. Sci. U.S.A.*, 84 (1987) 2727.
- [11] AMIT D. J., GUTFREUND H. and SOMPOLINSKY H., *Phys. Rev. A*, 35 (1987) 2293.
- [12] VAN KAMPEN N. G., *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam) 1981.
- [13] An equation of this kind has also been derived by PERETTO P. and NIEZ J. J., *IEEE Trans. Syst. Man Cybern.*, 16 (1986) 73.
- [14] BUHMANN J., Thesis, Technische Universität München, 1988.