INFORMATICA, 1996, Vol. 7, No. 2, 255–267

A MARKOVIAN STUDY OF RECURRENT NEURAL NETWORKS WITH STOCHASTIC DYNAMICS

Daniela ZAHARIE

Faculty of Mathematics, The West University of Timişoara
V. Pârvan bv. 4, 1900 Timişoara, Romania

Abstract. Recurrent neural networks of binary stochastic units with a general distribution function are studied using Markov chains theory. Sufficient conditions for ergodicity are established and under some assumptions, the stationary distribution is determined. The relation between fixed points and absorbing states is studied both theoretically and through simulations. For numerical studies the notion of almost absorbing state is introduced.

Key words: stochastic neural network, ergodicity, absorbing state.

1. Introduction. Neural networks with feedback connections, known as Little-Hopfield models, provide computing models that perform like auto-associative memories (Hopfield, 1982; Hertz *et al.*, 1991) or are capable to solve optimization problems (Tagliarini, 1991; Hertz *et al.*, 1991).

The stochastic modelling of neural networks is motivated by some characteristics of the biological neurons (random fluctuations of threshold, stochastic nature of synaptic processes etc.) and by improvements of retrieval capabilities of these systems when they are used like auto-associative memories (Bressloff and Taylor, 1990; Hertz *et al.*, 1991).

Reccurent neural networks of binary neurons with stochastic discrete dynamics were mainly studied, on the basis of their analogy with Ising spin glasses, through statistical mechanics techniques (Perreto, 1984; Amit *et al.*, 1985a; Amit *et al.*, 1985b; Derrida, 1990; Treves, 1990). An excellent review of the main results obtained in this field can be found in Hertz *et al.* (1991). A network of this type consists of N units, each of them receiving signals from all others (sometimes even from itselfs) and producing a binary output (for example ± 1) on the basis of a distribution probability (known as Glauber distribution) P_G:

$$P_G(y_i = +1) = \frac{1}{1 + \exp(-\beta(h_i - \theta_i))},$$
$$P_G(y_i = -1) = \frac{1}{1 + \exp(\beta(h_i - \theta_i))}. \quad (1)$$

In relations (1) y_i denotes the output produced by the i-th unit, h_i denotes the total input received by the same unit and θ_i denotes the neuronal threshold. Usually, $h_i = \sum_{j=1}^{N} w_{ij} y_j$ where $w_{ij} \in \mathbf{R}$ denotes the strength of the connection between the unit j and the unit i. The parameter β specifies the slope, in 0, of the distribution function.

This form of P_G created the opportunity to apply statistical mechanics tools in neural networks studies. For general probability distribution functions the analogy with spin glasses is no more available, thus the same technique cannot be applied.

On the other hand, under some assumptions, neural networks with stochastic discrete dynamics define Markov chains, so their behaviour could be studied using the theory of Markov processes. In fact, this tool was already used by Bressloff (Bressloff and Taylor., 1990) and François (François *et al.*, 1992) but only in the study of the above mentioned model with Glauber dynamics.

In this paper we shall study, using results of Markov chains theory, the synchronous discrete dynamics of a network with stochastic units characterized through a general probability distribution function defined by:

$$P(y_i = +1) = f(h_i),$$
$$P(y_i = -1) = 1 - f(h_i), \quad (2)$$

where $f : \mathbf{R} \to [0, 1]$ is a function on which we do not impose any restriction, at least for the moment.

The paper is organized as follows. In Section 2 the stochastic dynamics is presented, the transition probabilities are determined and some well known models are obtained like particular cases of the proposed model.

In Section 3 the ergodic properties of the system are studied and under some assumptions on f a general form of the stationary distribution function is obtained.

In Section 4 the relation between the absorbing states of the corresponding Markov chain and the fixed points of a related deterministic variant is studied.

In the same section the notion of almost absorbing state is introduced and some simulation results are presented.

In Section 5 some concluding remarks and open problems are presented.

2. Networks with synchronous stochastic dynamics. We shall consider a network of N stochastic units of type (2) characterized at a moment $t \in \{0, 1, \ldots\}$ through the output vector, $Y(t) = (y_1(t), y_2(t), \ldots, y_N(t)) \in \{-1, 1\}^N$, considered the state of the system. At one moment, all of the N units can change their state (i.e., the output they produce), thus the dynamics of the system is a synchronous (parallel) one (Little, 1974; Bressloff and Taylor, 1990).

In this case the evolution of the network's state can be described through:

$$\begin{aligned} P\big(y_i(t+1) = +1 | Y(t) = Y\big) &= f(\sum_{j=1}^{N} w_{ij} y_j - \theta_i), \\ P\big(y_i(t+1) = -1 | Y(t) = Y\big) &= 1 - f(\sum_{j=1}^{N} w_{ij} y_j - \theta_i), \end{aligned} \tag{3}$$

with $f : \mathbf{R} \to [0, 1]$ a function and $P(\cdot|\cdot)$ a conditional probability.

Clearly, the state at the moment $(t+1)$, $Y(t+1)$, depends only on the state at the moment t, $Y(t)$, thus $\mathcal{Y}_t = \{Y(t) | t \in \{0, 1, \ldots\}\}$ is an homogeneous Markov chain with 2^N states: $\{Y^1, Y^2, \ldots, Y^{2^N}\}$.

Let us denote through $P(k, l)$ the probability of one-step transition from state $Y^k = (y_1^k, \ldots, y_N^k)$ to state $Y^l = (y_1^l, \ldots, y_N^l)$: $P(k, l) = P(Y(t+1) = Y^l | Y(t) = Y^k)$. Using the fact that the dynamics of the network is synchronous, follows that, for each unit i, the output at moment $(t+1)$, $y_i(t+1)$, depends only on the outputs produced at the previous moment, thus on $y_1(t), \ldots, y_N(t)$. Hence

$$P(k, l) = \prod_{i=1}^{N} P\big(y_i(t+1) = y_i^l | Y(t) = Y^k\big).$$

On the other hand, because $y_i^l \in \{-1, 1\}$ relations (3) can be written in a single relation:

$$P\big(y_i(t+1) = y_i^l | Y(t) = Y^k\big) = \frac{1 - y_i^l}{2} + y_i^l f\Big(\sum_{j=1}^{N} w_{ij} y_j^k - \theta_i\Big).$$

Using these remarks, follows that the one-step transition probabilities, $P(k,l)$, can be expressed through:

$$P(k,l) = \prod_{i=1}^{N} \left(\frac{1-y_i^l}{2} + y_i^l f(\sum_{j=1}^{N} w_{ij} y_j^k - \theta_i) \right). \tag{4}$$

Sometimes it is useful to change f with $g : \mathbf{R} \to [-1,1]$, $g(u) = 2f(u) - 1$, for any $u \in \mathbf{R}$. In this case (4) becomes:

$$P(k,l) = \prod_{i=1}^{N} \frac{1 + y_i^l g(\sum_{j=1}^{N} w_{ij} y_j^k - \theta_i)}{2}. \tag{5}$$

In the following we shall use the notation $h_i^k = \sum_{j=1}^{N} w_{ij} y_j^k$. It is easy to verify that the matrix T with $T_{kl} = P(k,l)$, for $k,l \in \{1,\ldots,2^N\}$ is a stochastic matrix because:

$$\sum_{l=1}^{2^N} P(k,l) = \prod_{i=1}^{N} \left[f(h_i^k - \theta_i) + (1 - f(h_i^k - \theta_i)) \right] = 1.$$

A system with stochastic dynamics will, in general, evolve differently every time it is run. To obtain some informations of the system behaviour we can, for example, reduce them to a deterministic variant or study them for large values of t (the asymptotic behaviour).

There are at least two strategies to reduce the stochastic dynamics (3) to a deterministic one:

(a) At each moment the system will go in the state with the greatest probability: if $Y(t) = Y^k$ then $Y(t+1) = Y^l$ with $P(k,l) = \max_j P(k,j)$. It follows that $y_i^l = \mathrm{sgn}(g(h_i^k - \theta_i))$ with sgn the classical signum function ($\mathrm{sgn}(u) = -1, u < 0$; $\mathrm{sgn}(u) = 1, u \geqslant 0$). This dynamics includes, like a particular case, the deterministic Little model (Little, 1974).

(b) At each moment the system will go in a state which is obtained through averaging: if $Y(t) = Y^k$ then $y_i(t+1) = E(y_i^l | Y(t) = Y^k)$ with $E(\cdot|\cdot)$ the conditional mean of the binary random variable y_i^l. Thus $y_i(t+1) = g(h_i^k - \theta_i)$ which is the dynamics of a deterministic neural network with continuous units (Marcus and Westervelt, 1989).

On the other hand, any variant of the Little model with stochastic threshold (Bressloff and Taylor, 1990) can be obtained from the dynamics (3) if we take $f = F_\theta$, where F_θ is the distribution function of the random variable θ.

To conclude this section we remark that the network with stochastic dynamics (3) is a general one, which includes some of other known models. The asymptotic behaviour of the system described through (3) is studied in the next section.

3. Ergodic properties. It is known (Perreto, 1984; Bressloff and Taylor, 1990; Derrida, 1990; François *et al.*, 1992) that a network of stochastic units of type (1) and synchronous dynamics defines an ergodic Markov chain and the expression of the stationary distribution is known too. For the stochastic units with general distribution function (2) the main results which we obtained are presented in the next two propositions.

Proposition 3.1. *If* $f(\mathbf{R}) \subset (0,1)$ *(or equivalently* $g(\mathbf{R}) \subset (-1,1)$*) then* $\mathcal{Y}_t$ *is an ergodic Markov chain.*

Proof. From (4) and (5) and the hypothesis of the proposition follows that $P(k,l) \in (0,1)$ for all $k,l \in \{1,\ldots,2^N\}$, so T is a regular stochastic matrix. Applying the Perron-Frobenius theorem follows that $\mathcal{Y}_t$ is an ergodic Markov chain.

For an ergodic Markov chain the stationary distribution is unique but, in general, is difficult to determine its analytical expression. There are only some particular cases when the stationary distribution can be explicitly calculated.

For the general stochastic dynamics (3) we obtained:

Proposition 3.2. *If* $g(\mathbf{R}) \subset (-1,1)$ *and*

$$\sum_{k=1}^{2^N}\prod_{i=1}^{N}\sqrt{\frac{1+y_i^l g(h_i^k-\theta_i)}{1-y_i^l g(h_i^k-\theta_i)}} = \sum_{k=1}^{2^N}\prod_{i=1}^{N}\sqrt{\frac{1+y_i^k g(h_i^l-\theta_i)}{1-y_i^k g(h_i^l-\theta_i)}} \tag{6}$$

for all $Y^l, Y^k \in \{-1,1\}^N$ *then the stationary distribution is*

$$P^*(k) = \frac{1}{Z}\prod_{i=1}^{N}\frac{1}{\sqrt{1-g^2(h_i^k-\theta_i)}} \tag{7}$$

with

$$Z = \sum_{p=1}^{2^N}\prod_{i=1}^{N}\frac{1}{\sqrt{1-g^2(h_i^p-\theta_i)}}$$

a normalization constant.

Proof. The unicity of the stationary distribution of an ergodic Markov chain implies that it is sufficient to verify:

$$\sum_{k=1}^{2^N} P(k,l)P^*(k) = P^*(l)$$

for all $l \in \{1, \dots, 2^N\}$. This follows through some calculations:

$$\sum_{k=1}^{2^N} P(k,l)P^*(k) = \sum_{k=1}^{2^N} \frac{1}{2^N Z} \prod_{i=1}^{N} \frac{1 + y_i^l g(h_i^k - \theta_i)}{\sqrt{1 - g^2(h_i^k - \theta_i)}}$$
$$= \sum_{k=1}^{2^N} \frac{1}{2^N Z} \prod_{i=1}^{N} \frac{1 + y_i^l g(h_i^k - \theta_i)}{\sqrt{1 - (y_i^l)^2 g^2(h_i^k - \theta_i)}} = \sum_{k=1}^{2^N} \frac{1}{2^N Z} \prod_{i=1}^{N} \sqrt{\frac{1 + y_i^l g(h_i^k - \theta_i)}{1 - y_i^l g(h_i^k - \theta_i)}}$$
$$= \sum_{k=1}^{2^N} \frac{1}{2^N Z} \prod_{i=1}^{N} \sqrt{\frac{1 + y_i^k g(h_i^l - \theta_i)}{1 - y_i^k g(h_i^l - \theta_i)}}$$
$$= \sum_{k=1}^{2^N} P(l,k)P^*(l) = P^*(l).$$

The classical result for stochastic units of type (1) can be obtained like a particular case if we take $g(u) = \tanh(\beta u)$. Indeed, after some calculations follows:

$$\prod_{i=1}^{N} \frac{1 + y_i^l \tanh(\beta(h_i^k - \theta_i))}{1 - y_i^l \tanh(\beta(h_i^k - \theta_i))} = \exp\left(2\beta \sum_{i,j=1}^{N} w_{ij} y_i^l y_j^k\right).$$

If the connections between units are symmetric ($w_{ij} = w_{ji}$ for all $i, j \in \{1, \dots, N\}$) then $\exp(2\beta \sum_{i,j=1}^{N} w_{ij} y_i^l y_j^k) = \exp(2\beta \sum_{i,j=1}^{N} w_{ij} y_i^k y_j^l)$ thus the hypothesis of Proposition 3.2 is satisfied. So we reobtained for the case of a synchronous Glauber dynamics the stationary probability

$$P^*(k) = \frac{1}{Z} \prod_{i=1}^{N} \cosh(\beta(h_i^k - \theta_i))$$

with Z a normalization constant.

We must remark that for this particular case it is satisfied even a more restrictive condition than (6), the equality of corresponding terms:

$$\prod_{i=1}^{N}\sqrt{\frac{1+y_i^l g(h_i^k-\theta_i)}{1-y_i^l g(h_i^k-\theta_i)}}=\prod_{i=1}^{N}\sqrt{\frac{1+y_i^k g(h_i^l-\theta_i)}{1-y_i^k g(h_i^l-\theta_i)}}.$$

From this relation results that for the stochastic dynamics (1) it is satisfied the hypothesis of a known result (Bressloff and Taylor, 1990):

If $\{Y(t)|t\in \mathbf{N}\}$ *is an ergodic Markov chain with a finite set of states,* $\{Y^1,\ldots,Y^m\}$, *and if there exists a function* $a:\mathbf{R}\to\mathbf{R}_+^*$ *such that the transition probabilities satisfy:*

$$\frac{P(k,l)}{P(l,k)}=\frac{a(Y^l)}{a(Y^k)},$$

then the stationary distribution is

$$P^*(k)=\frac{a(Y^k)}{\sum\limits_{p=1}^{m}a(Y^p)};$$

for a particular choice of a, $a(u)=1/\sqrt{(1-g^2(u))}$.

In the general case when the relation (6) is satisfied but the corresponding terms are not equal, the Proposition 3.2. can be used but this last result cannot be applied. However, even the hypothesis of Proposition 3.2 could be too restrictive.

4. Absorbing states and fixed points. A deterministic neural network which implements an auto-associative memory for static vectors is characterized through the fact that the corresponding dynamical system has some fixed points which have the property of local asymptotic stability. These fixed points are included into the system through an appropriate choice of the parameters w_{ij}.

From a theoretical viewpoint the ergodicity is an undesirable property for networks used for auto-associative memory applications. There are at least two ways to breakdown the ergodicity (Bressloff and Taylor, 1990): considering the thermodynamic limit, $N\to\infty$, variant which allows in the case of dynamics (1) the use of statistical mechanics tools, or choosing $f(g)$ such that $f(\mathbf{R})\cap\{0,1\}\neq\varnothing$ ($g(\mathbf{R})\cap\{-1,1\}\neq\varnothing$) variant which we will study in the following.

The relation between the absorbing states Y^k (characterized through $P(k,k) = 1$) and the fixed points of a deterministic variant is presented in the next two propositions.

Proposition 4.1. *A vector $Y^k \in \{-1,1\}^N$ is an absorbing state for the Markov chain $\mathcal{Y}_t$ if and only if Y^k is a fixed point for $G : \{-1,1\}^N \to [-1,1]^N$ with $G(Y) = (g(h_1 - \theta_1), \ldots, g(h_N - \theta_N))^T$.*

Proof. Y^k is an absorbing state if and only if $P(k,k) = 1$ that implies

$$\frac{1 + y_i^k g(h_i^k - \theta_i)}{2} = 1$$

for any $i \in \{1, \ldots, N\}$. For $y_i^k \in \{-1,1\}$ this is equivalent with $y_i^k = g(h_i^k - \theta_i)$.

For a particular choice of the parameters w_{ij} which assures the storage of some vectors $\{\xi^1, \ldots, \xi^p\} \in \{-1,1\}^N$ (Hebb rule):

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \qquad i,j \in \{1, \ldots, N\}, \tag{8}$$

the absorbing states are characterized by:

Proposition 4.2. *If w_{ij} are established through (8) starting from a set of uncorrelated vectors $S = \{\xi^1, \ldots, \xi^p\}$ ($(\xi^\mu)^T \xi^\nu = 0$, for any $\mu \neq \nu$) then a vector $\xi^\nu \in S$ is an absorbing state if and only if*

$$g(\xi_i^\nu - \theta_i) = \xi_i^\nu \quad \text{for all} \quad i \in \{1, \ldots, N\}. \tag{9}$$

Proof. If S is a set of uncorrelated vectors then $\sum_{j=1}^N w_{ij} \xi_j^\nu = \xi_i^\nu$, so applying Proposition 4.1 follows that ξ^ν is an absorbing state if and only if $\xi_i^\nu = g(\xi_i^\nu - \theta_i)$, for any $i \in \{1, \ldots, N\}$.

For $\theta_i \in (-1,1)$, one of the simplest and most used function g which satisfies (9) for all uncorrelated stored vectors is the signum function (sgn). But in this case the dynamics becomes deterministic. To keep the stochastic character of the dynamics and breakdown the ergodicity property we can choose g such that $g(u) = -1$ for $u \in (-1-\delta, -1+\delta)$, $g(u) = 1$ for $u \in (1-\delta, 1+\delta)$ (with $\delta > 0$ a little real value) and $g(u) \in (-1,1)$ for other values of u.

On the other hand we can keep the ergodicity property and impose the existence of almost absorbing states defined by:

DEFINITION 4.1. A state Y^k is an almost absorbing state if there exists $\varepsilon \in (0,1)$ a little constant such that:

$$P(k,k) > 1-\varepsilon. \tag{10}$$

The existence of almost absorbing states can be achieved using functions g_β which depend on a parameter β and which satisfy: i) $g_\beta(\mathbf{R}) \subset (-1,1)$ and ii) $\lim_{\beta\to\infty} g_\beta(\mathbf{R}) = \{-1,0,1\}$. Clearly, for $\beta < \infty$ the Markov chain $\mathcal{Y}_t$ cannot have absorbing states but if β has a high value then there exists almost absorbing states.

In numerical simulations we can use almost absorbing states instead of absorbing states to identify the stored vectors, but it is necessary to determine values of β which assure the property of almost absorbing state for all stored vectors.

In the following we shall study two functions g_β which have the property $\lim_{\beta\to\infty} g_\beta(u) \in \{-1,0,1\}$ for any $u \in \mathbf{R}$:

$$g_1(u) = \frac{\exp(\beta u) - 1}{\exp(\beta u) + 1}, \tag{11}$$

$$g_2(u) = \begin{cases} 1 - \exp(-\beta u) & \text{if } u \geqslant 0, \\ \exp(\beta u) - 1 & \text{if } u < 0. \end{cases} \tag{12}$$

The function g_1 corresponds to the classical Glauber dynamics and g_2 is a symmetric variant obtained starting with McCulloch Pitts units with exponential distributed random threshold. For these functions we have determined lower bounds of β which assure, in the uncorrelated case, for all stored vectors (through (8)) the property of almost absorbing state (10).

These lower bounds are:

1. for g_1:

$$\beta > \max_i \theta_i + \ln \frac{\sqrt[N]{1-\varepsilon}}{1-\sqrt[N]{1-\varepsilon}} = K_1,$$

2. for g_2:

$$\beta > \max_i \theta_i + \ln \frac{1}{2(1-\sqrt[N]{1-\varepsilon})} = K_2.$$

Some numerical values of K_1 and K_2 corresponding to different values of absorption probability $1-\varepsilon$, for the case $\theta_i = 0$, $i \in \{1,\ldots,N\}$, are presented in Table 1.

Table 1. Dependence of lower bounds K_1 and K_2 on absorption probability $1-\varepsilon$

$1-\varepsilon$	K_1	K_2
0.5	2.5249	1.9088
0.6	2.8404	2.2040
0.7	3.2082	2.5547
0.8	3.6847	3.0163
0.9	4.4417	3.7062
0.99	6.7968	6.1047
0.999	9.1044	8.4113

In the case of correlated vectors, theoretical lower bounds for β cannot be determined so easy. In this case we made some simulations studies for a little network with $N = 9$ which determines a Markov chain with 512 states. Some ($p = 2, 3, 4, 5, 6$) correlated vectors were successively embedded into the parameters of the network and we calculated the absorption probability of states corresponding to the stored vectors. For the deterministic case with g =sgn if $p > 3$ (for $N = 9$ the critical capacity (Amit *et al.*, 1985b) is $p_c = N\alpha_c = 9 \cdot 0.138 = 1.242$) some spurios absorbing states appears.

For the stochastic case we denote through $P_{\min}$ the minimal absorption probability of stored vectors. If S is the set of stored vectors, a vector $Y^k \notin S$ will be considered spurious almost absorbing state if $P(k, k) > P_{\min}$. For the variants with g_1 and g_2 (for different values of $\beta \in \{3, \ldots, 10\}$ spurious almost absorbing states appears only for $p > 5$. The dependence of $P_{\min}$ on β and p is presented in Fig. 1 for g_1, respectively in Fig. 2 for g_2.

5. Conclusions. The reformulation in the terms of a Markov process of the stochastic dynamics of a neural network could be useful to obtain some informations even in the cases when statistical mechanics tools cannot be applied.

Using the absorption probabilities, informations on the fixed points of a deterministic corresponding network (with discrete or continuous units) can be obtained. On the other hand the transition probabilities could be used to obtain informations on the region of attraction of the fixed points.

There are, too, some unsolved problems: (1) to find other functions g which satisfy (6) or to relax this condition; (2) to establish, for the correlated case, the

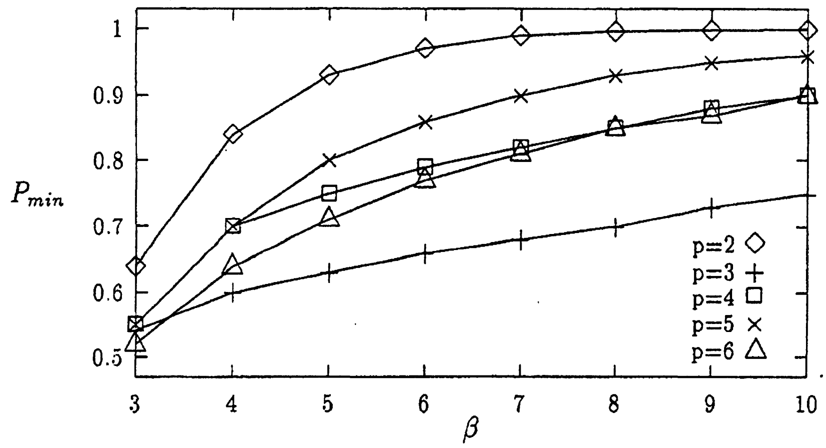

Fig. 1. Dependence of the minimal absorption probability for the stored vectors, $P_{\min}$, on the parameter β of g_1.

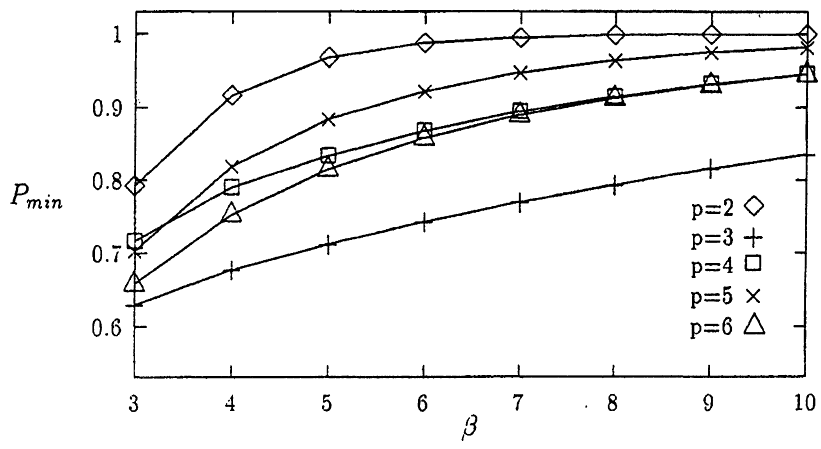

Fig. 2. Dependence of the minimal absorption probability for the stored vectors, $P_{\min}$, on the parameter β of g_2.

relation between the almost absorbing states and the fixed points of the deterministic dynamics; (3) to extend the Markovian study for stochastic networks of continuous valued units using results from the theory of Markov chains with infinite set of states.

REFERENCES

Amit, D.J., H. Gutfreund and H. Sompolinsky (1985). Spin-glass models of neural networks. *Physical Review A,* **32**(2), 1007–1018.

Amit, D.J., H.Gutfreund and H.Sompolinsky (1985b). Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters,* **55**(14), 1530–1533.

Bressloff, P.C., and J.G. Taylor (1990). Random iterative networks. *Physical Review A,* **41**(2), 1126-1137.

Derrida, B. (1990). Dynamical phase transitions in spin models and automata. In H. van Beijëren (Ed.), *Fundamental Problems in Statistical Mechanics,* Vol. 7. Elsevier Science Publisher B.V., pp. 273–309.

François, O., J. Demongeot, and T. Herve (1992). Convergence of a self-organizing stochastic neural network. *Neural Networks,* **5**, 277–282.

Hertz, J., A. Krogh and R.G. Palmer (1991). *Introduction to the Theory of Neural Computation.* Addison-Wesley P.C., 327 pp.

Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Science, U.S.A.,* **79**, 2554–2558.

Little, W.A. (1974). The existence of persistent states. *Mathematical Bioscience,* **19**, 101–120.

Marcus, C.M., and R.M. Westervelt (1989). Dynamics of iterated map neural networks. *Physical Review A,* **40**(1), 501–504.

Perreto, P. (1984). Collective properties of neural networks: a statistical physics approach. *Biological Cybernetics,* **50**, 51–62.

Tagliarini, G.A. (1991). Optimization using neural networks. *IEEE Transactions on Computers,* **40**(12), 1347–1358.

Treves, A. (1990). Graded-response neurons and information encodings in autoassociative memories. *Physical Review A,* **42**(4), 2418–2430.

Received January 1996

D. Zaharie born in 1965, graduated the Informatics section of the Faculty of Mathematics-Informatics from the University of Timişoara in 1987. At present she is a lecturer at the Department of Probability Theory and Applied Mathematics of the same faculty. Her research interests include neural networks, learning processes and stochastic modelling.

REKURENTINIŲ NEURO TINKLŲ SU STOCHASTINE DINAMIKA MARKOVINĖ ANALIZĖ

Daniela ZAHARIE

Rekurentiniai neuro tinklai, sudaryti iš stochastinių elementų su bendra pasiskirstymo funkcija, yra nagrinėjami, naudojantis Markovo grandinių teorija. Nustatytos pakankamos ergodiškumo sąlygos, o padarius tam tikras prielaidas, įvertintas stacionarus pasiskirstymas. Sąryšis tarp fiksuotų taškų ir absorbuojančių būklių yra nagrinėjamas teoriškai ir modeliavimo būdu. Eksperimentiniams tyrimams įvesta beveik absorbuojančios būklės sąvoka.