PHYSICAL REVIEW A VOLUME 43, NUMBER 12 15 JUNE 1991

Diluted-neural-network model with higher-order interactions

F. A. Tamarit,* D. A. Stariolo, and E. M. F. Curado
Centro Brasileiro de Pesquisas Fisicas, Rua Xavier Sigaud 150, 22290 Rio de Janeiro, Brazil

(Received 23 January 1991)

We study the retrieval properties of an extremely diluted neural network with multispin interactions. We calculate the phase diagram and show that a retrieval phase appears through a first-order phase transition provided the interactions are of order greater than 2.

In recent years a considerable effort in the modeling of neural networks has been devoted to improving the storage and retrieval properties of the networks and to making them biologically plausible. From the biological point of view, dilution and asymmetry have been shown to be two important elements, and have motivated a lot of work.[1–4] In particular, Derrida, Gardner, and Zippelius[1] have solved exactly the dynamics of strongly diluted asymmetric Hopfield model, obtaining an improvement of the storage capacity. Another important contribution has been the calculation of the number of metastable states in diluted Hopfield models carried out by Treves and Amit.[4] In another line of research, different authors have studied the effect of introducing multispin interactions in the models,[5–7] again finding an enhancement in the capacity of the net.

In this paper we join the two elements discussed above and study the dynamics and retrieval properties of a strongly diluted model of neural networks, with interactions between R neurons. We first consider the following Hamiltonian:

$$H=-\sum_{i_1,i_2,\ldots,i_R}^{N} J_{i_1 i_2 \cdots i_R} S_{i_1} S_{i_2} \cdots S_{i_R} \tag{1}$$

where the S_i's are N Ising variables that represent the states of the neurons and the synaptic interactions are given by a generalization of the Hebb rule,

$$J_{i_i i_2 \cdots i_R}=C_{i_1 i_2 \cdots i_R} \sum_{\mu=1}^{p} \xi_{i_1}^{\mu} \xi_{i_2}^{\mu} \cdots \xi_{i_R}^{\mu} \; . \tag{2}$$

Here the $\{\xi_i^\mu=\pm 1\}$ are p random uncorrelated patterns that we want to store and the $C_{i_1 i_2 \cdots i_R}$ are the dilution parameters. Next the $C_{i_1 i_2 \cdots i_R}$ are chosen randomly according to the following distribution:

$$\rho(C_{i_1 i_2 \cdots i_R})=\frac{C}{N^{R-1}}\delta(C_{i_1 i_2 \cdots i_R}-1) + \left[1-\frac{C}{N^{R-1}}\right]\delta(C_{i_1 i_2 \cdots i_R}) \; . \tag{3}$$

Note that this form of dilution destroys the symmetry of the couplings and the C is the mean number of plackets in the net.

The dynamics of the system is governed by a heat-bath Monte Carlo process for which all spins are updated simultaneously at time $t+1$ according to the following rule

$$S_i(t+1)=\begin{cases} 1 \quad \text{with probability}\{1+\exp[-2h_i(t)/T]\}^{-1} \\ -1 \quad \text{with probability}\{1+\exp[+2h_i(t)/T]\}^{-1} \end{cases} \tag{4}$$

where $h_i(t)$ is the local field at site i at time t:

$$h_i(t)=\sum_{i_2,\ldots i_R}^{N} C_{ii_2,\ldots,i_R} \sum_{\mu=1}^{P} \xi_i^\mu \xi_{i_2}^\mu \cdots \xi_{i_R}^\mu S_{i_2} \cdots S_{i_R} \tag{5}$$

and T is a generalized temperature that measures the noise level of the system.

We consider the evolution of a confirmation $\{S_i\}$ that has a macroscopic overlap with the pattern $\mu=1$ and a microscopic overlap with the other $p-1$ patterns. Following the ideas introduced by Derrida, Gardner, and Zippelius[1] we find a recurrence relation for $m(t)$ defined by

$$m(t)=\frac{1}{N}\sum_{i=1}^{N}\langle \xi_i^1 S_i(t)\rangle \tag{6}$$

of the form

$$m(t+1)=f(m(t)) \tag{7}$$

and then we study the retrieval properties of the model analyzing its asymptotic behavior. ($\langle\;\rangle$ means a thermal average.)

The local field at site i can be written as

$$h_i(t)=\xi_i^1\left[\sum_{\delta=1}^{K}\xi_{i_2}^1\cdots\xi_{i_R}^1 S_{i_2}\cdots S_{i_R} + \xi_i^1\sum_{\delta=1}^{K}\sum_{\mu>1}^{P}\xi_i^\mu\cdots\xi_{i_R}^\mu S_{i_2}\cdots S_{i_R}\right] \tag{8}$$

where $\delta=1,2,\ldots,K$ denotes a sum over the K plackets that contains the site i after dilution. Let s and n be the number of negative terms in the first and second sums of (8) respectively, then $h_i(t)$ can take the value $h_i(t)$

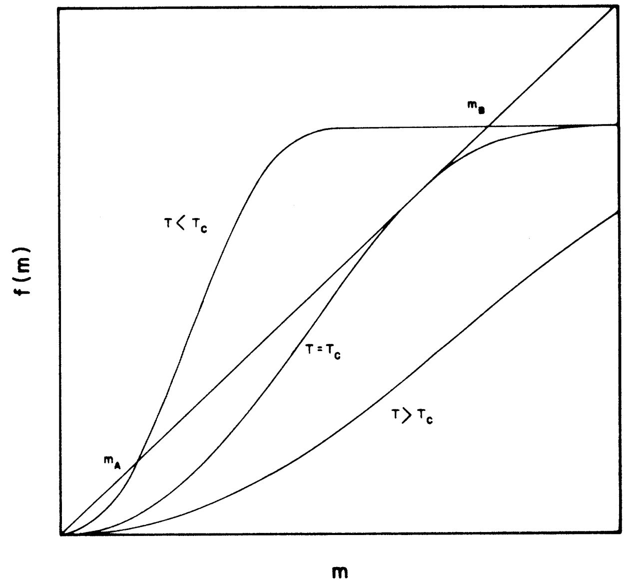

FIG. 1. Graphical solution of Eq. (13). At T_c the system undergoes a first-order phase transition. For $T>T_c$, $m=0$ is the unique fixed point, while for $T<T_c$ there are two stable fixed points, $m=0$ and m_B, and one unstable fixed point m_A.

$=KP-2(n+s)$ with probability

$$P(s,n)=\frac{1}{2^{K(p-1)}}\begin{bmatrix}K(P-1)\\ s\end{bmatrix}\begin{bmatrix}k\\ n\end{bmatrix}q^{K-n}(1-q)^n \tag{9}$$

where

$$g=\sum_{\substack{z=0\\ z\ \text{even}}}^{(R-1)}\begin{bmatrix}R-1\\ z\end{bmatrix}\left[\frac{1+m}{2}\right]^{R-z-1}\left[\frac{1-m}{2}\right]^{z} \tag{10}$$

with the sum running over the even values of z. Then after both a thermal and a disorder average we obtain

$$f(m)=\sum_{K=0}^{\infty}\frac{C^K e^{-c}}{K!}\sum_{n=0}^{K}\sum_{s=0}^{K(P-1)}\frac{q^{K-n}(1-q)^n}{2^{K(p-1)}}\begin{bmatrix}K\\ n\end{bmatrix}\times\begin{bmatrix}K(P-1)\\ s\end{bmatrix}\times\tanh\left[\frac{KP-2(s+n)}{T}\right]. \tag{11}$$

TABLE I. Critical storage capacity and the corresponding percentage of errors for different values of R at zero temperature.

R	α_c	% of errors
3	0.2525	7.65
4	0.1975	3.55
5	0.1725	2.25
6	0.1575	1.65
10	0.1285	0.65
100	0.0765	0.45
1000	0.0565	2.5×10^{-2}
10 000	0.0525	1.5×10^{-3}
100 000	0.052 25	5.0×10^{-4}

TABLE II. Critical temperature and the corresponding percentage of errors for different values of r for $\alpha=0$.

R	T_c	% of errors
3	0.5825	9.37
4	0.4955	6.17
5	0.4515	3.92
6	0.4215	0.51
10	0.3645	7.5×10^{-2}
100	0.2395	5.0×10^{-3}
1000	0.1825	5.0×10^{-4}
10 000	0.1485	5.0×10^{-5}
100 000	0.1255	5.0×10^{-6}

This result is valid provided that $(R-1)C \ll \ln N$.

Let α and T_0 be defined by

$$\alpha=\frac{p-1}{C},\quad T_0=\frac{T}{C}\ . \tag{12}$$

Next, we consider the limit C and $p\rightarrow\infty$ with finite α. Due to the absence of correlations between spins, we can replace the local field by a Gaussian variable obtaining

$$f(m)=\frac{1}{\sqrt{\pi}}\int dx\,\exp(-x^2)\tanh\left[\frac{m^{R-1}-x\sqrt{2\alpha}}{T_0}\right]. \tag{13}$$

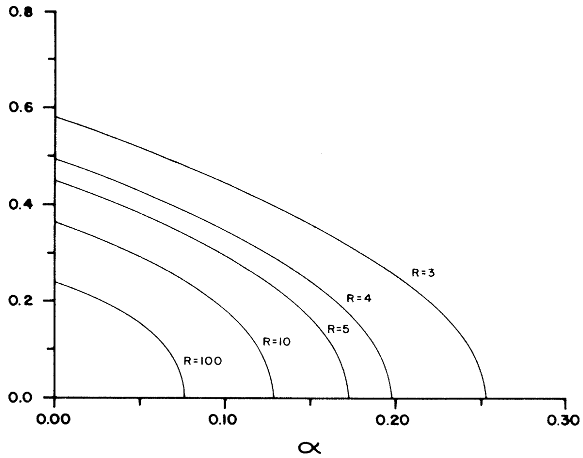

FIG. 2. Temperature vs storage capacity phase diagram for different orders of the interactions R. For each value of R we present the transition line, which separates the paramagnetic phase (for high values of T and α) from the retrieval one (low values).

TABLE III. Minimum value of the connectivity for which a retrieval solution exists for different numbers of stored patterns in the case $R=3$.

p	K_c
1	3
2	7
3	11
4	15
5	19
6	23

TABLE IV. Critical temperature below which a retrieval solution exists for different values of the connectivity K in the case $R=3$ and $p=2$.

K	T_c
7	1.59
8	2.31
9	2.97
10	3.60
11	4.25
12	4.85
20	9.57
40	21.00

The fixed points of the map (7) with f given by (13) and their stability will determine the retrieval properties of the model in this limit.

In what follows the analysis is restricted to the case $R>2$. The value $m=0$ is always a stable fixed point of the map. Moreover, for high temperature, it is the only solution; that means, the thermal noise avoids any possibility of retrieval. At a critical temperature T_c, for fixed R and α, the system undergoes a *first-order* transition to a low-temperature retrieval phase, unlike the case $R=2$ where the transition is continuous.[1] In this regime the map has two new fixed points m_A and m_B, as is schematically shown in Fig. 1, m_A being unstable and m_B stable for all temperatures lower than T_c. All initial conditions with value greater than m_A will converge to an overlap m_B and so it measures the size of the basin of attraction of the retrieval solution. The fixed point m_B is related to the percentage of errors in retrieval by the formula $(1-m_B)/2$.

For $T=0$, the relation (7) takes the following form:

$$m=\operatorname{erf}\left[\frac{m^{R-1}}{\sqrt{2\alpha}}\right]. \qquad (14)$$

In this limit, for fixed R, the system presents a retrieval phase for $0\leq\alpha\leq\alpha_c$ at which the system undergoes a transition to a nonretrieval phase. In Table I we show α_c and the corresponding percentage of errors for several values of R. From (14) one can demonstrate that for large values of R

$$m>1-\frac{1}{R}\left[\frac{1}{\pi\ln R}\right]^{1/2} \qquad (15)$$

provided that $\alpha<1/(2\ln R)$. In general, for any value of R and $\alpha\ll 1$, expanding (14) we obtain

$$m\simeq 1-\left[\frac{2\alpha}{\pi}\right]^{1/2}\exp\left[-\frac{1}{2\alpha}\right]. \qquad (16)$$

Note that this expression is independent of R. In particular, it agrees with the results obtained in the case $R=2$.[1]

In the limit $\alpha=0$ Eq. (7) takes the form

$$m=\tanh\left[\frac{m^{R-1}}{T}\right]. \qquad (17)$$

Table II shows the critical temperatures below which the system is in the retrieval phase, for several values of R, together with the percentage of errors at this temperature. In the large-R limit it can be shown that

$$m>1-\frac{2}{R}\ \text{ if } T<F\frac{2}{e^2\ln R}. \qquad (18)$$

For finite values of α and T the phase diagram is presented in Fig. 2 for several values of R. Note that the retrieval phase becomes smaller as the order of the interaction grows. As $R\to\infty$ both T_c and α_c go to zero logarithmically. The main characteristics of this phase diagram are similar to those found in the thermodynamics of multiconnected nondiluted models.[5,6]

Finally we analyze for $R=3$ the case in which only a finite number K of plackets contain a particular spin. Here f takes the following form:

$$f(m)=\sum_{n=0}^{K}\sum_{s=0}^{K(P-1)}\frac{q^{K-n}(1-q)^n}{2^{K(p-1)}}\begin{bmatrix}K\\ n\end{bmatrix}\begin{bmatrix}K(P-1)\\ s\end{bmatrix}\times\tanh\left[\frac{KP-2(s+n)}{T}\right]. \qquad (19)$$

Table III shows the minimal values K_c for which the system is able to store p patterns. For determining K_c we required a percentage of errors less than 5% in the retrieval. We note that in these conditions K_c grows arithmetically with p. In Table IV we present the critical temperature at which retrieval solution appears for the case $p=2$. We can see that it grows almost linearly with the connectivity. This result agrees with the one obtained for a diluted ferromagnet.[8]

In this paper we have studied the retrieval properties of an extremely diluted-neural-network model with multispin interactions. We have found that for any value of R a retrieval solution exists. Unlike the case $R=2$ we show that for $R\geq 3$ the transition to the retrieval phase is of first order. As the order of the interaction increases the values of the critical temperature and the critical storage capacity decrease.

---

*Also at Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba, Argentina.

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

[2] S. Bacci, G. Matto, and N. Parga, J. Phys. A **23**, 1801 (1990).

[3] D. J. Amit, in *Proceedings of Heidelberg Colloquium on Glassy Dynamics,* edited by J. L. van Hemmen and I. Morgenstern (Springer, Berlin, 1987), p. 430.

[4] A. Treves and D. J. Amit, J. Phys. A **21**, 3155 (1988).

[5] E. Gardner, J. Phys. A **20**, 3453 (1987).

[6] L. Abbot and Y. Arian, Phys. Rev. A **36**, 5091 (1987).

[7] D. Horn and M. Usher, J. Phys. (Paris) **49**, 389 (1988).

[8] O. Golinelli, Physica A **167**, 736 (1990).