GROUND STATES OF NEURAL NETWORK MODELS*

BY S. KOBE

Section of Physics, Technical University, Dresden**

AND A. SCHÜTTE

Section of Electronics, Humboldt-University, Berlin***

(Received October 3, 1988)

There is an analogy between the Ising spin glass model and the Hopfield model of a formal neural network. By choosing the strengths and the signs of the synaptic couplings it is possible to store p patterns as stable states of the network. The ground-state properties for systems up to N = 25 neurons are analyzed using algorithms of complex optimization. With increasing p also the number of other stable low-energy states grow up. Relations to the frustration in the network are discussed.

PACS numbers: 75.10.Hk, 87.30.-p

1. Introduction

Starting from the Ising spin glass model

$$E = - \sum_{i,j} I_{ij} S_i S_j$$

(1)

the analogy to the simple formal neural network model of Hopfield [1] can be realized by the following assumptions. The neurons can have two states characterized by the binary variable $S_i = \pm 1$. They are connected by the synaptic couplings $I_{ij}$ of different signs and strengths.

For general $I_{ij}$ the energy of the Ising model (1) as function of the states $S$ with the components $S_1, S_2, ..., S_N$ forms a complex energy landscape. The problem of finding the low-energy states of the system ("valleys") for given $I_{ij}$ is a task of combinatorical optimization [2]. For the neural network model first the inverse problem has to be solved.

---

* Presented at the 4th International Conference on Physics of Magnetic Materials, Szczyrk-Bila, Poland, September 4-10, 1988.
** Address: Sektion Physik, Technische Universität, DDR-8027 Dresden, Mommensnstr. 13, GDR.
*** Address: Sektion Elektronik, Humboldt-Universität, DDR-1040 Berlin, Invalidenstr. 110, GDR.
The synaptic couplings have to be chosen in such a way that the \( p \) given states \( S(k) \) of the system with \( k = 1, 2, \ldots, p \) become stable states with low energy values. The procedure of fitting the \( I_{ij} \) is called “learning”. After “learning” \( p \) patterns (prototypes) are stored in the network. Starting from a noise pattern a stored pattern can be retrieved by a simple relaxation procedure.

However, the retrieval process can be seriously disturbed, because besides the stored prototypes also so-called spurious states arise. Therefore the knowledge of all stable states with low energies are important for the valuation of the storage ability of the system. The purpose of this paper is to give an exact numerical analysis of the ground-state properties of networks with finite numbers \( N \) of neurons.

2. Model and method of calculation

Among the various “learning rules”, i.e. the rules to determine the coupling strengths \( I_{ij} \), the projection rule \([3, 4]\) has the advantage that all prototypes are global minima of the network. The matrix of prototypes \( \Sigma \) is formed by the column vectors \( S^{(1)} \ldots S^{(p)} \). Then the matrix of the synaptic couplings \( I \) with the elements \( I_{ij} \) is obtained by

\[
I = \Sigma \cdot \Sigma^+,
\]

where \( \Sigma^+ \) is the Moore–Penrose pseudo-inverse of \( \Sigma \) \([3]\). Dimensionless energy units are used. Following Kanter and Sompolinsky \([5]\) we set all diagonal terms \( I_{ij} \) equal zero.

We have investigated small systems with \( N = 16, 20 \) and \( 25 \) neurons for different numbers \( p \) of stored patterns, which can be also correlated ones. The problem of finding the ground state belongs to the class of NP-complete problems of the complex optimization \([6]\). Complete enumeration of all states is performed for systems up to \( N = 20 \), whereas for \( N = 25 \) the method “branch-and-bound” \([7]\) is used for an exact determination of all low-lying states.

The stability of a given state \( S \) against one-spin flips is checked by the condition \( S_i h_i \geq 0 \) for all \( i = 1, 2, \ldots, N \) of the components \( S_i \) of \( S \), where \( h_i = \Sigma I_{ij} S_j \) are the corresponding local fields.

3. Results and discussion

Fig. 1 shows the density of states of a system with \( N = 20 \) for different numbers of stored patterns \( p \). The storage of only one prototype (upper part of the figure) is possible as a single ground state which is well separated from all other states. In this case the system is non-frustrated. Embedding of more prototypes occurs at the expense of increasing frustration. The following consequences can be observed. The ground-state energy increases. Besides the prototypes also other spurious states become ground states of the network. Further spurious states arise as stable states with energies above the ground-state energy.

The gap structure of the density of states as well as the gap between the ground state and the other stable states disappear. In the case of \( z = p/N = 0.75 \) (15 prototypes, see the lower part of the figure) only a small part of all ground states represents the stored patterns.
Fig. 1. Histogram for the density of states \( n \) (in logarithmic scale) vs \( E \) for \( N = 20 \) and \( \alpha = p/N = 0.05 \) (upper part), 0.25 (middle part) and 0.75 (lower part) \( n \) = number of states in the energy interval \( \Delta E = 0.25 \); Stable states are shaded.

Fig. 2. Ratio (number of prototypes \( p \))/(number of ground states \( n_{GS} \)) vs \( \alpha = p/N \) for \( N = 16 \) (\( \times \)) and \( N = 25 \) (\( \bigcirc \)).
and, therefore, the network ceases to provide associative memory. This behaviour is also reflected in Fig. 2. It shows that the region of $\alpha$, for which a recall of memory without errors is possible, becomes smaller for increasing $N$.

The influence of the stable states with energies above the ground-state energy is shown in Fig. 3. These states disturb the retrieval process already for such $\alpha$, at which the prototypes are still the single ground states of the network.

In [8] a misfit parameter $m = 1 - E_{GS}/\Sigma |t_{ij}|$ is introduced as a quantitative measure for the frustration of an Ising system, where $E_{GS}$ is the ground-state energy. $m$ is independent
of the details of the model and the used learning rule. It can be also estimated for larger systems. The increase of the function \( m(x) \) (Fig. 4) suggests that \( m \) can be used as a parameter for the retrieval probability of a network.

In this paper properties of finite neural network models were analyzed by methods of the complex optimization. It is shown that the number of stored prototypes influences the frustration of the system. On the one hand, frustration is necessary for storage of more than one prototype, but its increase also restricts the retrieval behaviour of the considered model as an associative memory.

REFERENCES