Capacities of multiconnected memory models
D. Horn, M. Usher

To cite this version:

HAL Id: jpa-00210709
https://hal.archives-ouvertes.fr/jpa-00210709
Submitted on 1 Jan 1988

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Capacities of multiconnected memory models

D. Horn and M. Usher

School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel

(Reçu le 3 septembre 1987, accepté le 12 novembre 1987)

Abstract. — We study generalizations of the Hopfield model for associative memory which contain interactions of \( R \) spins with one another and allow for different weights for input patterns. Using probabilistic considerations we show that stability criteria lead to capacities which increase like powers of \( N^{R-1} \). Investigating numerically the basins of attraction we find behaviour which agrees with theoretical expectations. We introduce the more stringent definition of « coverage-capacity » by requiring the whole phase-space to be covered by the basins of attraction of the input patterns. Even under these conditions we find large numbers of patterns which can be used to design an output spectrum by varying the input weights.

1. Introduction.

The Hopfield model [1] is defined by the Hamiltonian

\[
H = - \sum_{ij} J_{ij} S_i S_j \quad J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_{i\mu} \xi_{j\mu} \quad (1)
\]

whose dynamical degrees of freedom are \( N \) classical spin variables which interact with one another through the couplings \( J_{ij} \). The binary vectors \( \xi_{i\mu} \) are interpreted as input patterns into the model, and they are supposed to become the fixed points to which the spin variables will flow under some relaxation method for this statistical-mechanics system. An example of dynamical equations of motion at zero-temperature is the time sequence defined by

\[
S_i(t + 1) = \text{sign} \left( \sum_{j} J_{ij} S_j(t) \right) \quad (2)
\]

where the updating is done for one spin at a time. Such models have a limited capacity [2] i.e. there is a limited number of input patterns \( p = \alpha N \) after which they lose their basins of attraction in the thermodynamic limit.

Several authors have recently investigated generalizations of the Hopfield model which include higher polynomial interactions representing multiconnected neural networks [3]. These can serve as practical models for biological systems or artificial-devices [4] and they have highly increased capacities [5-7]. We consider variations of such models defined by the Hamiltonian

\[
H = - N \sum_{\mu=1}^{p} \epsilon_\mu (m^\mu)^R \quad m^\mu = \frac{1}{N} \sum_{i=1}^{N} \xi_{i\mu} S_i \quad (3)
\]

where we introduce weights which can differentiate between the various patterns in our problem. Such generalizations of the \( R = 2 \) model were recently discussed in the literature [8]. In section 2 we use probabilistic arguments to study limiting values for the capacities as defined by the requirement of the existence of some basin of attraction in the neighbourhood of any original pattern. This can be extended to any number of weights and we look in particular into the case where we have only two different weights in our system. Section 3 is devoted to numerical investigations of...
the basins of attraction of the patterns. We check the relevance of the asymptotic result derived in section 2 to finite \( N \) systems, and show that the variation of the basins of attraction gives further support to this approach. We find that for a large range of \( p \) there is no interference of spurious minima, which leads us to introduce the concept of « coverage capacity » in section 4. This corresponds to the requirement that the input patterns cover the whole phase-space with their basins of attraction. We study it numerically for equal weights, and go on to demonstrate the possibility of designing a full spectrum of output states by using the freedom of attributing weights to the input patterns. We find in our examples that it is useful not to vary the weights by more than a factor of two, otherwise the probability of retrieval of the lower weights becomes very small.

2. Stability criteria.

The model of equation (3) is defined in terms of the order-parameters \[ m^\mu \] which represent the overlap of a spin configuration with a given pattern. Rewriting the same expression in terms of the independent spins \( S_i \) on the site \( i \) this Hamiltonian takes the form

\[
H = - \sum_{i_1 \ldots i_R} J_{i_1 \ldots i_R} S_{i_1} \ldots S_{i_R}
\]

with the couplings chosen as

\[
J_{i_1 \ldots i_R} = \frac{1}{N^{R-1}} \sum_{\mu = 1}^p \epsilon_{i_1}^\mu \ldots \epsilon_{i_R}^\mu .
\]

This is the model which we investigate here. We will also refer to a variant of this model in which the summation of equation (5) does not allow any two \( i \)-indices to be the same :

\[
H' = - \sum_{i_1 \neq i_2 \ldots \neq i_R} J_{i_1 \ldots i_R} S_{i_1} \ldots S_{i_R}.
\]

We will show that \( H' \) has a larger capacity than \( H \). Nonetheless we will use \( H \) in our numerical investigations because it is much easier to carry them out by employing the form (3).

The storage capacity of the model can be studied by testing the stability of the original patterns. We will employ a simple probabilistic approach to investigate this problem for \( R \gg 2 \). As we will see it allows us to obtain in a straightforward manner results which were previously obtained using the replica symmetry method [6]. We will start with the case in which all weights are equal, \( \epsilon_{i_\mu} = 1 \), following a well-known argumentation [9], which relies on representing the action of the Hamiltonian \( H \) on some spin \( S_i \) in terms of a local field \( h_i \). To leading order in \( N \), this field is given by

\[
h_i = R \sum_\mu \xi_i^\mu (m^\mu)^{R-1}.
\]

A stable solution to the serial dynamics of the system is one in which the spin aligns with the local field,

\[
S_i = \text{sign} (h_i).
\]

Looking for the stability of a configuration of spins \( S_i \) which coincides with the pattern \( \xi \) we try to satisfy equation (8) in a self-consistent way. First we note that in this case \( m^\mu = 1 \) whereas all other \( m^\mu \) fluctuate around 0 with a standard deviation of \( 1/\sqrt{N} \). Let us separate these two kinds of contributions into signal and noise :

\[
h_i = R (\xi_i^\mu + \delta_i)
\]

\[
\delta_i = \sum_{\mu \neq \nu} \xi_i^\mu \xi_j^\nu - \frac{1}{N}.\]

The noise \( \delta \) vanishes on the average. Its standard deviation is calculated by looking at the statistical average of \( \delta^2 \) which leads to a finite result through all possible contractions of pairs \( \xi^\mu \). This takes the form

\[
\sigma^2 = \frac{pN^{1-R} \gamma}{2^{R-1}(R-1)!}.
\]

where \( \gamma \) is the combinatorial weight

\[
\gamma = \frac{(2R - 2)!}{2^{R-1}(R-1)!},
\]

which takes into account all possible pairing of indices needed for the contraction.

It is clear from equation (10) that the natural choice of the maximal number of retrievable patterns should be represented as

\[
p = \frac{\alpha N^{R-1}}{\gamma}
\]

which coincides with the standard notation [2] for \( R = 2 \) where \( \gamma = 1 \). \( \alpha \) is determined by the requirement that the noise \( \delta \) does not reverse the sign of the signal in equation (9) and, therefore, equation (8) holds. The probability \( P \) that this will indeed be the case is

\[
P = 1 + \text{erf} \left( \frac{1}{\sqrt{2}} \right).
\]

This leads to the conclusion that

\[
\alpha = \frac{1}{2c \log N}
\]

where \( c = 1 \) if one requires the stability of a single.
pattern \( (p^N = O(1)) \) and \( c = R \) if one demands perfect retrieval of all patterns \( (p^{pN} = O(1)) \).

An analogous consideration for the Hamiltonian \( H' \) leads to a smaller combinatorial weight because of the smaller number of possible contractions:

\[
\gamma' = (R - 1)! \tag{15}
\]

Replacing \( \gamma \) in equation (12) by this smaller factor we find that the corresponding capacities turn out to be bigger. We conclude that \( H' \) has a larger capacity than \( H \), yet both follow the same functional behaviour.

The argument of exact retrieval was employed by Peretto and Niez [4] for finite values of \( N \). Gardner, as well as Abbot and Arian [6] have used the thermodynamic limit and continued to consider imperfect retrieval by employing the replica symmetry approach. It is known in the usual case of \( R = 2 \) that requiring a finite overlap \( m \) instead of perfect retrieval \( (m = 1) \) one ends up with finite \( \alpha \) in the thermodynamic limit [2, 9] instead of the vanishing value of (14). This remains true [6] for \( R > 2 \) and may be simply shown [7] by modifying equation (9) accordingly to represent the local field for the \( N(1 + m)/2 \) retrievable spins:

\[
h_i = R(m^{R-1} \xi^v + \delta_i) \tag{16}
\]

The main assumption of our probabilistic approach is that \( \delta \) continues to lead to the same results as before. In other words, in spite of the fact that the \( S_i \) of the stable configuration are not identical with \( \xi^v \) we assume that they are independent of the other \( \xi^p \), thus leading to equation (10) and (11) in the same way that \( \delta \) of equation (9) led to them. The intuitive reason for this assumption is that since the number of input patterns is much larger than \( N \) they cover in an over-complete fashion all directions and, therefore, do not introduce a net influence which distorts that of the nearest vector (*). This is not the case for \( R = 2 \) where it is known that the bias of all other vectors has to be taken into account. The probability that equation (8) holds takes now the form

\[
P = \frac{1 + \text{erf} \left( \frac{m^{R-1}/ \sqrt{2} \alpha}{2} \right)}{2} \tag{17}
\]

Let us find the probability that some configuration of spins with overlap \( m \) with the pattern \( \xi^v \) becomes stable. This is given by the sum over all spin configurations which have this overlap, weighted by the probabilities that each one of the retrievable spins has the same value as the pattern (i.e. the noise does not reverse the signal of the pattern) while the opposite happens for each non-retrievable spin. This leads to the expression

\[
\left( \frac{N}{N(1 + m)/2(1 - P)^{N(1 - m)/2}} \right)
\]

which is maximized by the choice \( P = (1 + m)/2 \) meaning that the probability of right retrieval in equation (17) is also the relative average number of spins which agree with the original pattern. This amounts to an equation for the overlap

\[
m = \text{erf} \left( \frac{m^{R-1}/ \sqrt{2} \alpha}{2} \right) \tag{18}
\]

which has solutions for only a limited range of \( \alpha \). In the case \( R = 3 \) we find that \( \alpha < \alpha_{cr} = 0.25 \). For this \( \alpha \) the average overlap \( m \) is 0.84. \( \alpha_{cr} \) decreases slightly with increasing \( R \). For \( R = 4 \) we find that \( \alpha_{cr} = 0.195 \). The corresponding overlap increases to 0.92. The graphical solutions of equation (18) for these two limiting cases are displayed in figure 1.

\[ \text{Fig. 1.} \quad \text{Graphical solutions of equation (18). The curves represent both sides of the equations for the cases } \alpha = 0.25, \ R = 3 \text{ and } \alpha = 0.195, \ R = 4 \text{ which are the critical } \alpha \text{ values.} \]

Equation (18) coincides with the condition obtained from a replica-symmetry analysis [6], thus justifying our simple assumption that made it possible to apply probabilistic considerations searching for stable minima in the neighbourhood of an original pattern. Our analysis can be easily generalized to the case of different weights. Let us assume, for simplicity, that all patterns in equation (3) are assigned one of two possible weights which are chosen as \( \varepsilon_1 > \varepsilon_2 \). Correspondingly we will have two different \( \alpha_1 \) and \( \alpha_2 \) describing the capacities of the two families of patterns. Requiring stability of a pattern from the first family one obtains equations similar to the ones above with the replacement

\[
\alpha \rightarrow \alpha_1 + \left( \frac{\varepsilon_2}{\varepsilon_1} \right)^2 \alpha_2.
\]

(*) We thank H. Sompolinsky for a discussion of this point.
Allowing finite overlaps we are led to the two conditions
\[ \alpha_1 + \left( \frac{e_2}{e_1} \right)^2 \alpha_2 < \alpha_{cr} \]  
(19a)
\[ \alpha_2 + \left( \frac{e_1}{e_2} \right)^2 \alpha_1 < \alpha_{cr} \]  
(19b)

where \( \alpha_{cr} \) is the same one as before. If the second constraint is obeyed so will be the first one. This means that if the situation is such that patterns of type 2 are stable then also patterns of type 1 will be stable. The opposite is however not true. It is quite possible to have a situation in which patterns of type 1 are stable but patterns of type 2 are not. This is natural since the patterns of type 1 lie deeper than those of type 2. The gap between the two constraints widens as the ratio \( \frac{e_1}{e_2} \) increases, as demonstrated in figure 2. Clearly even if the patterns of type 2 are unstable their mere appearance in the Hamiltonian limits the capacity of patterns of type 1 since the former add to the destabilizing noise. This effect diminishes as \( e_2 \) decreases, as can also be seen in figure 2.


The stability criteria of the previous section find the point at which the basins of attraction vanish. To investigate the latter with more detail one has to rely on numerical methods. The first question we have to settle is how the asymptotic results of the thermodynamic limit, such as equation (18), relate to numerical calculations at finite \( N \). We have run Monte-Carlo programs for the case \( R = 3 \) and a series of \( N \) spins, \( N \) varying in the range of 50 to 200. In each case we chose \( \alpha N^{2/3} \) random input patterns to construct the model, and then measured the probability that a spin-configuration which starts out identical to an input-pattern, flows into a fixed-point which has overlap \( m \) with the original pattern. The results are displayed in the histograms of figure 3.

Fig. 2. — The constraints of equations (19a) and (19b) are represented for \( e_2/e_1 = 0.8 \) and 0.6 by the full lines and the dashed lines respectively. Equation (19b) is the stronger constraint. Even if it is not obeyed, the presence of patterns of type 2 have an effect on the remaining constraint on \( \alpha_1 \).

Fig. 3. — Histograms of the probability that a spin-configuration corresponding to an initial pattern flows into a fixed-point which has overlap \( m \) with its origin. Shown are two values of \( \alpha \), above and below \( \alpha_{cr} \), for \( R = 3 \) and \( N = 50 \) and 200.
For $\alpha > \alpha_{cr}$ one expects the distribution to shift toward lower $m$ values as $N$ increases while for $\alpha < \alpha_{cr}$ it should concentrate in the high $m$ region. In this way one builds up a critical behaviour at infinite $N$. The different trends are clearly seen for the two $\alpha$-values displayed in figure 3. We find stability to occur around $\alpha = 0.25$, in accordance with our previous estimate. Thus for 100 spins we can store under these conditions about 800 patterns. This is a marked improvement over the 10 or so which are allowed for $R = 2$. Moreover, since $R$ is odd, the Hamiltonian is no longer symmetric under the sign inversion $\xi \rightarrow -\xi$ as is the case for even $R$, hence one never retrieves the inverse pattern.

The algorithm which one uses for testing stability starts with a trial state which is equal to one of the input patterns and checks the fixed point to which it flows. This may be regarded as insufficient for any practical purposes which a device, built on these principles, may serve. We do expect the model to allow the flow into the pattern from a certain region in phase-space close to the pattern. We are interested therefore in some additional information about the size of the basins of attraction of the input patterns. Following conventional procedures [10] we measure it numerically by starting with an arbitrary input pattern, inverting randomly its bits with probability $d/N$ (where $d$ denotes the Hamming distance) and choosing all the resulting spin-configurations which have overlaps bigger than

$$m = 1 - \frac{2d}{N}$$

with the original pattern. We let all these spin-configurations flow into their fixed points and determine whether they fall in the neighbourhood of the original pattern. To get a feeling about the structure of the basins of attraction we ask for the $m$ value above which the probability of flowing into the original pattern is larger than 0.9. Clearly this value will change both with $p$, the number of input patterns, and with $R$, the power of the interaction. The results for $R = 3$ and 4 are the data points shown in figure 4.

It is quite evident from figure 4 that the basins of attraction widen considerably as the power of the interaction increases from 3 to 4, since lower $m$ means larger Hamming distances over which the attraction of the input pattern is felt. Note however that for $R = 4$ there exists the global symmetry of spin-inversion, i.e. together with every pattern one finds also its inverse as a source of attraction. In the results displayed in this figure we took this doubling into account by using here $p$ to denote all stable patterns. Using the theoretical considerations of equation (16) we can explain the nature of the curves shown here. Since the signal varies like $m^{R-1}$ and the noise is proportional to $(p/N^{R-1})^{1/2}$ we expect these curves to display a behaviour of

$$m \propto p^{\frac{1}{2(R-1)} N^{-1/2}}.$$  \hspace{1cm} (20)

This behaviour is borne out by the data, as seen from the fits which vary like $p^{16}$ and $p^{16}$ for $R = 3$ and 4 respectively. These fits describe well the data over quite a large range of $p$. This range is also characterized by the fact that almost all initial spin-configurations flow into one of the original patterns. In other words, the spin-configurations do not end up in a spurious state which corresponds to some other local minimum of the energy. When this power behaviour fails we reach a region where the spurious states have a dominant effect. In the next section we will investigate further the region where the spurious states are negligible.


In the Hopfield model one finds that only a minute fraction of the $2^N$ states which constitute the phase-space is covered by the total sum of the basins of attraction of the original patterns. The situation improves considerably as one increases $R$, as long as the number of patterns remains moderate. The higher $R$ models enable us therefore to develop a new notion which we call coverage-capacity $C$, by which we mean the maximal number of patterns which allows for the complete coverage of the phase-space. We envisage the division of phase-space into $C$ domains ($2C$ domains for even $R$) in which all states flow into one of the original $C$ patterns (or its inverse for even $R$). To turn this into a measurable...
quantity we have to specify margins of error and use a probabilistic definition. In our algorithm we will choose a random state of spins and require it to flow into a state which has an overlap of $|m| \geq M$ with one of the original patterns with probability $P$. In the results presented in figure 5 we use $M = 1$ and $P = 0.96$.

Figure 5 displays the coverage-capacity of the $R = 3$ and 4 models described by $H$ of equation (3) in which all weights were assumed equal. Due to our stringent requirements the resulting numbers are much smaller than the capacities determined by the stability considerations. Within this range they rise with $N$ but slower than (12). Nonetheless the values of $C$ are still very impressive. For a model with several tens of spins we are able to specify hundreds of ground-states chosen at will with a very small margin of error (i.e. small number of spurious minima). Allowing ourselves the use of the interaction of $R$ spins with one another we can construct spin models for almost any set of patterns which we wish to store and reproduce successfully.

![Figure 5](image_url)

**Fig. 5.** Coverage-capacities for the $R = 3$ and $R = 4$ models as determined by numerical investigations in which we required an arbitrary pattern to flow into one of the original patterns with probability $P = 0.96$.

As long as the number of patterns is smaller than $C$ we can in fact construct a whole spectrum of output states with relative ease by introducing different weights into the original Hamiltonian. This is displayed in table I where we look into models based on $N = 50$ spins and $p = 200$ input patterns. These input patterns are generated randomly and are assigned one of four possible weights. We have therefore four groups of 50 input patterns with weights of 1, 0.8, 0.6 and 0.4. As one may expect [8], the higher weights dominate over the lower ones. The probability to flow into a pattern diminishes considerably as its weight decreases. Nonetheless the output spectrum reflects well the input energies.

<table>
<thead>
<tr>
<th>$R$</th>
<th>$\langle -E/N \rangle$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.17</td>
<td>0.65</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>0.49</td>
</tr>
<tr>
<td>6</td>
<td>1.02</td>
<td>0.45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\varepsilon$</th>
<th>1</th>
<th>0.8</th>
<th>0.6</th>
<th>0.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R = 4$</td>
<td>1.17</td>
<td>0.98</td>
<td>0.79</td>
<td>0.58</td>
</tr>
<tr>
<td>$P$</td>
<td>0.65</td>
<td>0.27</td>
<td>0.07</td>
<td>0.005</td>
</tr>
<tr>
<td>$R = 5$</td>
<td>1.00</td>
<td>0.80</td>
<td>0.60</td>
<td>0.40</td>
</tr>
<tr>
<td>$P$</td>
<td>0.49</td>
<td>0.33</td>
<td>0.14</td>
<td>0.04</td>
</tr>
<tr>
<td>$R = 6$</td>
<td>1.02</td>
<td>0.82</td>
<td>0.62</td>
<td>0.42</td>
</tr>
<tr>
<td>$P$</td>
<td>0.45</td>
<td>0.32</td>
<td>0.17</td>
<td>0.06</td>
</tr>
</tbody>
</table>

For even $R$ the average energy is being shifted by an amount of

$$\delta E = \frac{1}{N^{R/2}} \binom{R}{2} \sum \varepsilon \mu$$

which is the expectation value of a random spin-configuration for this Hamiltonian. Odd $R$ should show no shift on the average. In both cases one expects of course fluctuations which can be estimated in the same fashion, leading to

$$\sigma^2 = \frac{1}{N^R} \binom{2R}{2} \sum \varepsilon \mu.$$  

This approach can be further refined to produce a custom-designed output spectrum. Starting from a set of input patterns with designated energies which are used as input weights, one produces a certain output spectrum. In general it will have small variations in the energies, as well as some unwanted states. These can be corrected by adjusting accordingly the input weights. Spurious states which appear in the output can be discarded by including them in the input with corresponding negative weights. A similar suggestion was made in the past [11] for the Hopfield model. Here it works perfectly. We have tried it at the edge of the coverage-capacity and succeeded to increase considerably the probability of flowing into the desired patterns after a small number of iterations.

Finally we wish to present the variation of the basins of attraction as a function of the weight. This we do by considering two groups of random input patterns which are assigned the two weights

$$\varepsilon_1 = 1, \quad \varepsilon_2 = 1 - x$$

and measuring the corresponding probabilities of flowing into them from any random start. Choosing an overall number which is smaller than $C$ we have essentially $P_1 = P_2 = \frac{1}{2}$ for $x = 0$, whereas as
$x \to 1$ we expect $P_1 = 1$ and $P_2 = 0$. Some examples of this behaviour are shown in figure 6. All of them display a linear behaviour in $x$ over the range $0 < x < 0.5$ after which $P_2$ becomes very small. Both this figure and table I lead to the conclusion that it is advisable to use moderate ratios of weights, preferably less than 2, in order to avoid running into low probabilities for the lower weights.

Fig. 6. — Probabilities of flowing into two families of patterns with weights $\epsilon_1 = 1$ $\epsilon_2 = 1 - x$ are shown for the cases $R = 3$ $p = 50$ (□), $R = 4$ $p = 200$ (×), $R = 5$ $p = 400$ (○) which are all below coverage-capacity of $N = 50$ spins.

References