Pattern recognition in Hopfield type networks with a finite range of connections
Eva Koscielny-Bunde

To cite this version:
Short Communication

Pattern recognition in Hopfield type networks with a finite range of connections

Eva Koscielny-Bunde(1)

(1) Fachbereich Informatik, Universität Hamburg, D-2000 Hamburg 50, F.R.G.

(Received on June 15, 1990, accepted on June 20, 1990)

Abstract. — We study pattern recognition in linear Hopfield type networks of \( N \) neurons where each neuron is connected to the \( z \) subsequent neurons such that the state of the \( i^{th} \) neuron at time \( t + 1 \) is determined by the states of neurons \( i + 1, \ldots, i + z \) at time \( t \). We find that for small values of \( z/N \) the retrieval behavior differs considerably from the behavior of diluted Hopfield networks. The maximum number of random patterns that can be retrieved increases in a non-linear way with \( z \) and the asymptotic mean overlap between input and output patterns decreases sharply as \( z \) is decreased and reaches zero at a finite value of \( z \).

In recent years, Hopfield neural networks [1] have been studied extensively (for recent reviews see e.g. [2-4]). The Hopfield network provides a standard for neural nets since its long time retrieval behavior could be treated analytically [5] by solving for the equilibrium statistical mechanics of the net. The solution could be achieved since all neurons in the net are interconnected.

In this paper we report on numerical studies of linear Hopfield type nets of (usually \( N = 400 \)) neurons where each neuron is connected only to a certain number of other neurons, and periodic boundary conditions are employed. Previous work in this direction was mainly devoted to nearest neighbor and next nearest neighbor connections, or to dilute (damaged) networks where a certain fraction of connections was randomly cut [6-10].

The network consists of \( N \) neurons, each of them can be in two states \( S_i = \pm 1 \). The coupling strength \( J_{i,j} \) between pairs of connected neurons is determined by the \( M \) random patterns \( \{ \xi^\mu \} \equiv (\xi_1^\mu, \ldots, \xi_N^\mu) \), \( \mu = 1, 2, \ldots, M \), which are to be stored in the network:

\[
J_{i,j} \equiv J_{j,i} = \frac{1}{N} \sum_{\mu=1}^{M} \xi_i^\mu \xi_j^\mu. \tag{1}
\]

The state of neuron \( i \) at a given time \( t + 1 \) is obtained from the states of the \( z \) subsequent neurons at time \( t \) by

\[
S_i(t + 1) = \text{sign} \left[ \sum_{j=i+1}^{i+z} J_{i,j} S_j(t) \right]. \tag{2}
\]
For $z = N - 1$, (2) is identical to the retrieval process in Hopfield nets where each neuron receives input signals from all other neurons and sends output signals to all other neurons. For $z < N$, each neuron $i$ receives input signals only from the $z$ subsequent neurons $i + 1, \ldots, i + z$. The set of neurons influenced by neuron $i$ and the set of neurons influencing neuron $i$ partially overlap for $z \geq N/2$. For $z < N/2$, the neurons in both sets are all different.

A particular pattern $\{\xi^\nu\}$ is called stored by the network, if an initial configuration $\{S(0)\} = (S_1(0), \ldots, S_N(0))$ near $\{\xi^\nu\}$ develops under the dynamic rule (2) to a state whose overlap

$$q^\nu = \frac{1}{N} \sum_{i=1}^{N} S_i \xi_i^\nu,$$

with $\{\xi^\nu\}$ is large.

The maximum number of patterns $M_c$ that can be stored depends on the size of the network, the difference between input and output patterns (initial noise) and the number of connections $z$ per neuron. According to Amit et al. [5], we have in the limit of large Hopfield networks ($z = N - 1 \rightarrow \infty$),

$$M_c \approx 0.14 N, \quad (z = N - 1)$$

In this work, we study how equation (4) is modified for $z < N - 1$, and how the retrieval rate decays in the nonretrieval regime.

In the numerical study, first $M$ random patterns $\{\xi^\mu\}$ are generated and the connection strengths $J_{i,j}$ are calculated. Then one of the nominated patterns is chosen and a noisy input pattern is generated where the state of each neuron is changed with probability $p_n$. The dynamic rule (2) is applied and the time dependent overlap $q^\mu(t)$ is calculated. The run stops when $q^\mu(t)$ stays constant for six time steps or oscillations of period 2 with constant amplitude occur. This procedure is repeated for typically $10^3$ initial configurations, which are chosen from different nominated patterns. By averaging the results we obtained, for a given number $M$ of patterns and fixed initial noise $p_n$, the mean retrieval overlap $\langle q \rangle$ as a function of $z$.

Figure 1 shows the dependence of $\langle q \rangle$ on $z/N$, for pure input patterns ($p_n = 0$) and networks with $N = 120$, 200 and 400 neurons. In the different networks, the ratio $M/N$ has been kept constant at 0.075. One can distinguish between 3 régimes: (1) the retrieval regime for large $z$ where $M/z > 0.14$ and $\langle q \rangle$ is close to one; (2) an intermediate transient regime where $\langle q \rangle$ is well below one but greater than zero; (3) a disordered régime below some threshold value $z_c$, where $\langle q \rangle = 0$ and the net behaves as if all connections were cut; $z_c$ seems to increase slightly with increasing size of the network. The width of the transient region (2) shrinks if the size of the network is increased.

Figure 2 shows, again for pure input patterns, the variation of $\langle q \rangle$ with the number of nominated patterns $M$, for $N = 400$ fixed and three values of $z$: $z = 399$, 240, and 80; $z = 399$ corresponds to the conventional Hopfield net. The result is plotted versus $M/z$ rather than $M/N$. For $M/z$ not too large, the data approximately collapse to a single line. At intermediate values of $M/z$, the data split into three curves. For $z = N - 1$, a strong decrease in the mean overlap is followed by some smooth behavior where $\langle q \rangle$ varies only little with $M/z$ and seems to reach a plateau $> 0$ at large values of $M/z$ (see also [5, 11]). For $z = 240$, again the curve drops sharply and then decreases less strongly with increasing $M/z$. For $z = 80$, finally, the mean overlap shows a sharp decrease with $M/z$ and reaches zero at a finite value of $M/z$ roughly at $M/z \approx 0.3$.

We have compared this behavior with the case where the connections between neurons have been randomly destroyed [6, 10]. In this case, $z$ can be referred to as mean number of connections per neuron which is related to the concentration $p$ of damaged bonds by $z = (1 - p)N$. In figure 2, the results for the random damage are indicated by black triangles and squares, describing $p = 0.4$.
Fig. 1. — Plot of the asymptotic mean overlap as a function of $z/N$ for $N = 120(\triangle)$, $200(\square)$, and $400(\circ)$. The number of stored patterns per neuron is kept fixed, $M/N = 0.075$, and $p_n = 0$.

and $p = 0.8$, respectively. The curves differ strongly from the corresponding results for $z/N = 0.6$ and 0.2. In the diluted network, the mean retrieval overlap does not scale with $M/z$ and varies comparatively weakly with $M/z$.

Figure 3 finally shows the maximum number of patterns per connecting bond $M_c/z$ that can be stored with less than 0.5 percent error, as a function of $z$ for three values of the initial noise
For pure input patterns \( (p_n = 0) \) and \( z/N > 1/4 \), \( M_c/z \) is approximately the same as for the Hopfield net,
\[
M_c(z)/z \cong M_c(N)/N \sim 0.14,
\]
which also agrees with the situation in diluted Hopfield nets [10]. Below \( z/N \cong 1/4 \), \( M_c(z)/z \) decreases almost linearly with \( z/N \), with the exception \( M_c/z = 1 \) for \( M = 1 \). Almost the same curve is obtained for \( p_n = 0.2 \) except for \( M = 1 \) and at large values of \( z/N \) where \( M_c/z \) is considerably smaller than for pure input patterns. For very noisy input patterns \( (p_n = 0.4) \), the curve again is rather flat and except at small values of \( z/N \) depends rather weakly on \( z \). It is interesting to note that in all cases \( M_c/z \) does not show any pronounced features around the "critical" value \( z/N = 1/2 \) where the sets of input and output neurons start to have no overlap.

The behavior shown in figure 3 differs clearly from the behavior in diluted networks where \( M_c/z \) is described by a simple crossover behavior. Below some crossover value \( z_c (p_n) \) which decreases with increasing noise level \( p_n \), \( M_c/z \) is roughly constant and follows (5), while above \( z_c M_c/z \) bends down [10]. In particular for large noise, \( M_c/z \) decreases as \( 1/z \) for \( z \) approaching one, which is very different from the behavior observed here.

**Fig. 3.** — Maximum fraction \( M_c/z \) of learned random patterns in a network of 400 neurons that are recognized with less than 0.5 percent error, as a function of the number of connections per neuron \( z \), for the initial noise \( p_n = 0(\square), 0.2(\circ), \) and \( 0.4(\triangle) \).

It is a pleasure for me to thank Greg Kohring and Dietrich Stauffer for valuable discussions and critical reading of the manuscript.

**References**

   Phys. 58 (1990) 1257 and in: Parallel Processing in Neural Systems and Computers, R. Eckmiller,
   G. Hartmann, G. Hauske Eds. (North Holland, Amsterdam) 1990, p. 495.

Cet article a été imprimé avec le Macro Package "Editions de Physique Avril 1990".