Spatial asymmetric retrieval states in binary attractor neural network

Kostadin Koroutchev[*,†] and Elka Korutcheva[**,‡]

[*]*Escuela Politecnica Superior, Universidad Autónoma de Madrid, Spain*
[†]*Inst. for Personal Computing and Communication Systems, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria*
[**]*Depto. de Fisica Fundamental, UNED, Madrid, Spain*
[‡]*G.Nadjakov Inst. Solid State Physics, Bulgarian Academy of Sciences, 1784 Sofia, Bulgaria*

Abstract. In this paper we show that during the retrieval process in a binary Hebb attractor neural network, spatial localized states can be observed when the connectivity of the network is distance-dependent and there is an asymmetry between the retrieval and the learning states.

INTRODUCTION

In a very recent publication [1] it was shown that using linear-threshold model neurons, Hebb learning rule, sparse coding and distance-dependent asymmetric connectivity, spatial asymmetric retrieval states (SAS) can be observed. Similar results have been reported in the case of Hebb binary model for associative neural network [2]. This asymmetric states are characterized by a spatial localization of the activity of the neurons, described by the formation of local bumps.

In this paper we show that the major factor to observe spatial asymmetric activity is the distance-dependent connectivity of the neurons in the network and the different level of activity in the retrieval and learning states.

Our model consists of an attractor neural network (NN) model of Hebbian type, formed by N binary neurons $\{S_i\}, S_i \in \{-1,1\}, i = 1,...,N$, storing p binary patterns $\eta_i^\mu, \mu \in \{1...p\}$, and we assume a symmetric connectivity between the neurons $c_{ij} = c_{ji} \in \{0,1\}, c_{ii} = 0$. $c_{ij} = 1$ means that neurons i and j are connected. We regard only connectivities in which the fluctuations between the individual connectivity are small, e.g. $\forall_i \sum_j c_{ij} \approx cN$, where c is the mean connectivity.

The learned patterns are drawn from the distribution:

$$P(\eta_i^\mu) = \frac{1+a}{2}\delta(\eta_i^\mu - 1) + \frac{1-a}{2}\delta(\eta_i^\mu + 1),$$

where the parameter a is the sparsity of the code. Further on we will work in terms of variables $\xi_i^\mu = \eta_i^\mu + a$.

To impose a condition on the mean activity, in order to introduce asymmetry between the learning and the retrieval states, we add an extra term H_a to the Hamiltonian

$$H_a = NR \sum_i S_i / N.$$

This term actually favors states with lower total activity $\sum_i S_i$ that is equivalent to decrease the number of active neurons, creating asymmetry between the leaning and the retrieval states. When $H_a = 0$ we have equal mean activity of the learned pattern and the the retrieval state.

ANALYTICAL ANALYSIS

For the analytical analysis of the SAS states, we consider the decomposition of the connectivity matrix c_{ij} by its eigenvectors $a_i^{(k)}$:

$$c_{ij} = \sum_k \lambda_k a_i^{(k)} a_j^{(k)}, \quad \sum_i a_i^{(k)} a_i^{(l)} = \delta_{kl}, \tag{1}$$

where λ_k are the corresponding (positive) eigenvalues.

Following the classical analysis of Amit et al. [3], we study binary Hopfield model [4]

$$H = -\frac{1}{cN} \sum_{ij\mu} S_i \xi_i^\mu c_{ij} \xi_j^\mu S_j - \sum_{v=1}^s h^v \sum_i \xi_i^v S_i + NR \overline{S_i b_i^0}, \tag{2}$$

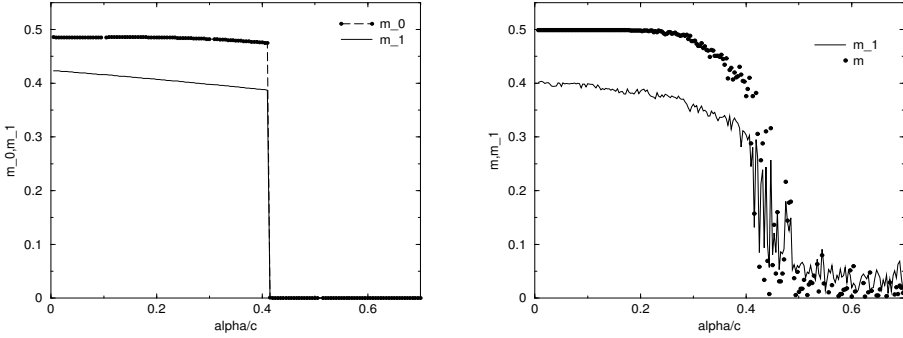where we have assumed Hebb's rule of learning [5].

By using the "replica formalism" [6], for the averaged free energy per neuron we get:

$$f = \lim_{n \to 0} \lim_{N \to \infty} \frac{-1}{\beta nN} (\langle\langle Z^n \rangle\rangle - 1), \tag{3}$$

where $\langle\langle...\rangle\rangle$ stands for the average over the pattern distribution $P(\xi_i^\mu)$, n is the number of the replicas, which later are taken to zero and β is the inverse temperature.

The saddle-point method [3], gives the following final form for the free energy per neuron:

$$
\begin{aligned}
f \quad = \quad & \frac{1}{2c}\alpha + \frac{1}{2}\sum_k (m_k)^2 + \frac{\alpha\beta(1-a^2)}{2}\sum_k r_k q_k + \frac{\alpha\beta(1-a^2)}{2}\sum_k \mu_k^2 r_k + \\
+ \quad & \frac{\alpha}{2\beta}\sum_k [\ln(1-\beta(1-a^2)\mu_k + \beta(1-a^2)q_k) - \\
- \quad & \beta(1-a^2)q_k(1-\beta(1-a^2)\mu_k + \beta(1-a^2)q_k)^{-1}] - \\
- \quad & \frac{1}{\beta}\int \frac{dz e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \ln 2\cosh\beta \left(z\sqrt{\alpha(1-a^2)\sum_l r_l b_i^l b_i^l} + \sum_l m_l \xi_i b_i^l + Rb_i^0 \right).
\end{aligned}
\tag{4}
$$

FIGURE 1. Left: The order parameteres m_0, m_1 for $a = 0.2$ and $R = 0.57$. The result of the simulations is represented on the right side.

In the last expression we have introduced the variables $\mu_k = \lambda_k/cN$ and $b_i^k \equiv a_i^{(k)}\sqrt{\lambda_k/c} \sim 1$ and we have used the fact that the average over a finite number of patters ξ^ν can be self-averaged [3].

The different order parameters, present in the above expression are the following: the overlap between patterns and neurons $m_k = \overline{(\eta - a)b_i^k S_i}$, where the k-mode corresponds to the k-th eigenvalue of the connectivity matrix and the over line is the space average, the neural activity, $q_k = \overline{\sum_i (b_i^k)^2 S_i^2}$ and r_k is the parameter conjugate to q_k. Finally, $\alpha = p/N$ is the storage capacity of the network.

RESULTS AND DISCUSSION

The analysis of the equations for the corresponding order parameters at zero temperature is shown in Fig. 1 (left). The sharp bound is due to the limitation of the equations only to the first two terms of the OP m, r, C, as well as to the finite scale effects.

In order to compare the results, we also performed computer simulations. To this aim we chose the network's topology to be a circular ring, with a distance measure

$$|i - j| \equiv \min(i - j + N \bmod N, j - i + N \bmod N)$$

and used the same connectivity as in Ref.[1] with typical connectivity distance $\sigma_x N$:

$$P(c_{ij} = 1) = c\left[\frac{1}{\sqrt{2\pi}\sigma_x N}e^{-(|i-j|/N)^2/2\sigma_x^2} + p_0\right].$$

The corresponding behavior of the order parameters m_0, m_1 in the zero-temperature case, obtained by the simulations, is represented in Fig. 1 (right). As can be seen, there is a good correspondence between the numerical solution of the analytical results and the results obtained by simulation.

The behavior of the information capacity, presented in the left panel of Fig.2, is non trivial and shows a well pronounced maximum for intermediate values of the sparsity of
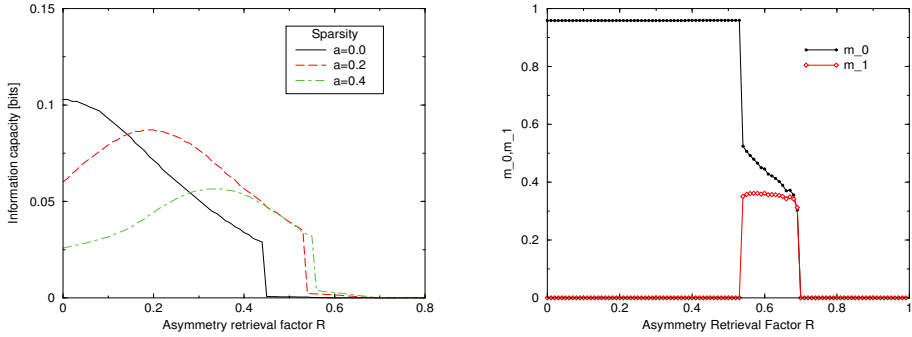
FIGURE 2. Left: the information capacity versus the asymmetry retrieval factor R for different values of the sparsity a. Right: m_0, m_1 for $a = 0.2$.

the code a. There is no maximum of the mutual information for $a = 0$. The corresponding behavior of the order parameters m_0 and m_1, presented in the right panel of Fig. 2, shows that inside the SAS region, their values are significantly different form zero, but the information capacity is very low although non zero.

The analysis of the corresponding phase diagram [7], shows that there exist states with $a = 0$ and spatial asymmetry retrieval states. Therefore, the sparsity of the code is not a necessary condition for SAS, but it makes the effect of SAS better pronounced.

On the other hand there is no state with SAS and $H_a = 0$, and therefore the asymmetry between the retrieval activity and the learned patterns is essential for the observation of the phenomenon.

In conclusion, we have shown that the presence of the term H_a, which introduces asymmetry between retrieval and learning states is sufficient for the existence of the spatially asymmetric retrieval states. The drop of the information capacity of the network is very significant in the phase transition point between symmetric and asymmetric retrieval states.

REFERENCES

1.   Y.Roudi and A.Treves, *JSTAT* P07010 (2004).
2.   K.Koroutchev and E.Korutcheva, *Preprint ICTP*, Trieste, Italy, IC/2004/91.
3.   D.Amit, H.Gutfreund and H.Sompolinsky, *Ann.Phys.*, **173**, 30 (1987).
4.   J.Hopfield, *Proc.Natl.Acad.Sci.USA*, **79** 2554 (1982).
5.   D.Hebb, *The Organization of Behavior: A Neurophysiological Theory*, Wiley, New York, 1949.
6.   M.Mézard, G.Parisi and M.-A.Virasoro, *Spin-glass theory and beyond*, World Scientific, 1987.
7.   K.Koroutchev and E.Korutcheva, Central European Journal of Physics, 2005, in press.