Artificial Neural Networks – ICANN 2006: 16th International Conference, Athens, Greece, September 10-14, 2006. Proceedings, Part I. Lecture Notes in Computer Science, Volumen 4131. Springer, 2006. 234-243.

DOI: http://dx.doi.org/10.1007/11840817_25

Improved Storage Capacity of Hebbian Learning Attractor Neural Network with Bump Formations

Kostadin Koroutchev * and Elka Korutcheva **

EPS, Universidad Autónoma de Madrid,
Cantoblanco, Madrid, 28049, Spain
k.koroutchev@uam.es

Depto. de Física Fundamental,
Universidad Nacional de Educación a Distancia,
c/ Senda del Rey 9,28080 Madrid, Spain
elka@fisfun.uned.es

Abstract. Recently, bump formations in attractor neural networks with distance dependent connectivities has become of increasing interest for investigation in the field of biological and computational neuroscience. Although the distance dependent connectivity is common in biological networks, a common fault of these network is the sharp drop of the number of patterns p that can remembered, when the activity changes from global to bump-like, than effectively makes these networks low effective. In this paper we represent a bump-based recursive network specially designed in order to increase its capacity, which is comparable with that of randomly connected sparse network. To this aim, we have tested a selection of 700 natural images on a network with $N = 64K$ neurons with connectivity per neuron C. We have shown that the capacity of the network is of order of C, that is in accordance with the capacity of highly diluted network. Preserving the number of connections per neuron, a non-trivial behavior with the radius of the connectivity has been observed. Our results show that the decrement of the capacity of the bumpy network can be avoided.

1 Introduction

Recently, the bump formations in recurrent neural networks have been analyzed in several investigations concerning linear-threshold units [1, 2], binary units [3] and Small-World networks of Integrate and Fire neurons [4]. These bump formations represent geometrically localized patterns of activity and have a size proportional to the number of connections per neuron.

* Also with: ICCS, Bulgarian Academy of Science.
** Also with: ISSP, Bulgarian Academy of Science.

It has been shown that the localized retrieval is due to the short-range connectivity of the networks and it could explain the behavior in structures of biological relevance as the neocortex [5].

In the case of linear-threshold neural network model, the signal-to-noise analysis has been used [1, 2] in the case of spatially organized networks. It has been shown that the retrieval states of the connected network have non-uniform activity profiles, when the connections are short-range enough, and that the level of localization increases by increasing the gain or the neuron's saturation level [2].

The investigation of the spontaneous activity bumps in Small-World networks (SW) [6, 7] of Integrate-and Fire neurons [4], has recently shown that the network retrieves when its connectivity is close to the random and displays localized bumps of activity, when the connectivity is close to the ordered. The two regimes are mutually exclusive in the range of the parameter governing the proportion of the long-range connections on the SW topology of Integrate-and-Fire network, while the two behaviors coexist in the case of linear-threshold and smoothly saturated units.

The result related to the appearance of bump formations have been recently reported by us [3] in the case of binary model for associative network. We demonstrated that the spatially asymmetric retrieval states (SAS) can be observed when the network is constrained to have a different activity compared to that induced by the patterns.

The model we studied in Ref.[3] has a symmetric and distance dependent connectivity for all neurons within an attractor neural network (NN) of Hebbian type formed by N binary neurons $\{S_i\}, S_i \in \{-1, 1\}, i = 1, ..., N$, storing p binary patterns $\eta_i^\mu, \mu \in \{1...p\}$ with symmetric connectivity between the neurons $c_{ij} = c_{ji} \in \{0, 1\}, c_{ii} = 0$.

The corresponding Hopfield model is [8]:

$$H = \frac{1}{N} \sum_{ij} J_{ij} S_i S_j, \tag{1}$$

with Hebbian learning rule:

$$J_{ij} = \frac{1}{c} \sum_{\mu=1}^{p} c_{ij} (\eta_i^\mu - a)(\eta_j^\mu - a). \tag{2}$$

The learned patterns are drawn from the following 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 [9]. The number of connections per neuron is $C \equiv cN$.

In order to introduce an asymmetry between the retrieval and the learning states, a condition on the mean activity of the network has to be imposed:

$$H_a = NR(\sum_i S_i/N - a).$$

Here the parameter R controls the affinity of the system toward the appearance of bump formations. This term favors states with lower total activity $\sum_i S_i$ that is equivalent to a decrease of the number of active neurons. Thus, the above term creates an asymmetry between the learning and the retrieval states. We showed in Ref.[3] that this condition is a necessary and a sufficient condition for the observation of bump formations. Similar observations have been reported in the case of linear-threshold network [1], where in order to observe bump formations, one has to constrain the activity of the network. The same is true in the case of smoothly saturating binary network [2], when the highest activity level, that can achieved by the neurons, is above the maximum activity of the units in the stored pattern.

As we have shown in Ref.[3], when the bump appears, the capacity of the network drops dramatically because the network decreases its effective size to the size of the bump.

On the other side, the spatially restricted activity means that the effective coding is restricted to very sparse coding. Usually, the capacity of the network, keeping patterns with sparsity a, is proportional to $1/a|\log a|$ and increases with the sparsity. Therefore a question arise: Is it possible to use the sparsity of the code, imposed by the bump in order to increase the capacity of the network?

In this paper we are trying to use explicitly the sparseness of the code, imposed by the bump appearance, in order to increase the capacity of a two-dimensional network that stores natural images. By means of simulations, we show that the capacity of the network can be increased to the limits typical for sparsely connected network, e.g. to achieve capacities of order of C.

2 Theoretical analysis

The theoretical analysis of the present model has been explained in details in Ref.[3]. Here we briefly present the main points of this analysis and the most important results.

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

$$c_{ij} = \sum_k b_i^{(k)} b_j^{(k)}, \tag{3}$$

with

$$b_i^k \equiv a_i^{(k)} \sqrt{\lambda_k / c} \tag{4}$$

where λ_k label the corresponding (positive) eigenvalues.

Following the classical analysis of Amit et al. [10], we study the following binary Hopfield model [8]:

$$H = -\frac{1}{cN} \sum_{ij\mu} S_i \xi_i^\mu c_{ij} \xi_j^\mu S_j - \sum_{\nu=1}^{s} h^\nu \sum_i \xi_i^\nu S_i + NR\overline{S_i b_i^0}, \tag{5}$$

where the Hamiltonian (5) has been represented in terms of the variables $\xi_i^\mu = \eta_i^\mu - a$. The second term in the Hamiltonian introduces a small external field h, which will tend later to zero. This term is necessary to take into account the finite numbers of overlaps that condense macroscopically. The third term imposes an asymmetry in the neural network's states, which is controlled by the value of the parameter R. This term is responsible for the bump formations. In Eq. (5) the over line means a spatial averaging $\overline{(.i)} = \frac{1}{N}\sum_i(.)$.

Following the classical analysis [10], we use the variables $m^\mu_{\rho k}$ for each replica ρ and each eigenvalue, which are the usual overlaps between neurons and patterns. In addition, we also use two other order parameters (OP)

$$q_k^{\rho,\sigma} = \overline{(b_i^k)^2 S_i^\rho S_i^\sigma},$$

which is related to the neuron activity, and the parameter $r_k^{\rho,\sigma}$, conjugate to

$q_k^{\rho,\sigma}$, that measures the noise in the system. The indexes $\rho, \sigma = 1, ..., n$ label the different replicas, while the role of the connectivity matrix is taken into account by the parameters b_i^k.

Finally, by using the replica symmetry ansatz and the saddle-point method [10], we obtain the following expression for the free energy per neuron:

$$\begin{aligned} f = {} & \frac{1}{2c}\alpha(1-a^2) + \frac{1}{2}\sum_k (m_k)^2 - \frac{\alpha\beta(1-a^2)}{2}\sum_k r_k q_k + \frac{\alpha\beta(1-a^2)}{2}\sum_k \mu_k r_k + \\ & + \frac{\alpha}{2\beta}\sum_k [\ln(1-\beta(1-a^2)\mu_k + \beta(1-a^2)q_k) - \\ & - \beta(1-a^2)q_k(1-\beta(1-a^2)\mu_k + \beta(1-a^2)q_k)^{-1}] - \\ & - \frac{1}{\beta}\int \frac{dz e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \overline{\ln 2\cosh\beta\left(z\sqrt{\alpha(1-a^2)\sum_l r_l b_i^l b_i^l} + \sum_l m_l \xi_i b_i^l + R b_i^0\right)}, \end{aligned} \tag{6}$$

where $\alpha = p/N$ is the storage capacity, $\mu_k = \lambda_k/cN$ and we have used the fact that the average over a finite number of patters ξ^ν can be self-averaged [3], [10].

The equations for the OP r_k, m_k and q_k are respectively:

$$r_k = \frac{q_k(1-a^2)}{(1-\beta(1-a^2)(\mu_k - q_k))^2}, \tag{7}$$

$$m_k = \int \frac{dz e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \overline{\xi_i b_i^k \tanh\beta\left(z\sqrt{\alpha(1-a^2)\sum_l r_l b_i^l b_i^l} + \sum_l m_l \xi_i b_i^l + R b_i^0\right)} \tag{8}$$

and

$$q_k = \int \frac{dz e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \overline{(b_i^k)^2 \tanh^2\beta\left(z\sqrt{\alpha(1-a^2)\sum_l r_l b_i^l b_i^l} + \sum_l m_l \xi_i b_i^l + R b_i^0\right)}. \tag{9}$$

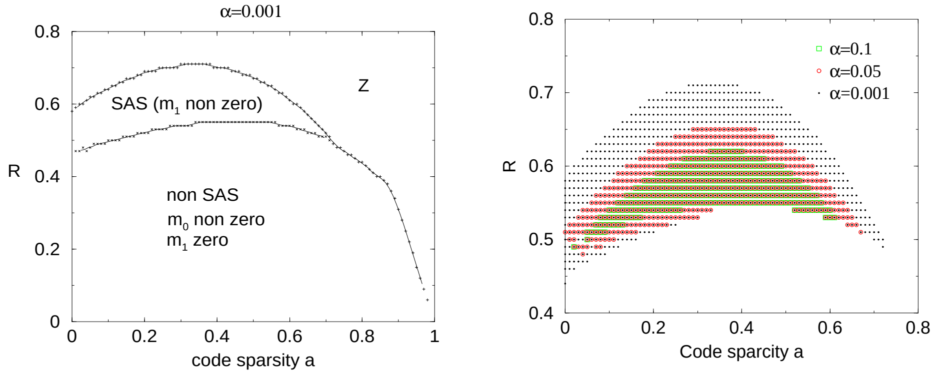

Fig. 1. Left:Phase diagram for $\alpha = 0.001$ ($\alpha/c = 0.02$). The SAS region, where local bumps are observed, is relatively large. The Z region corresponds to trivial solutions for the overlap. Right: SAS region for different values of α. High values of α limit the SAS region.

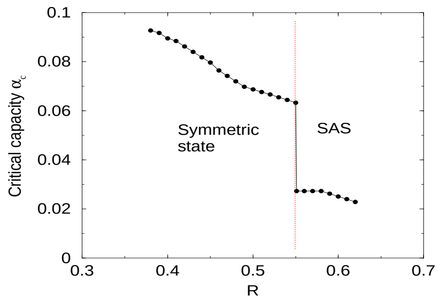

Fig. 2. The critical storage capacity α_c as a function of R for $a = 0.4$, showing a drop on the transition to a bump state.

The numerical analysis of the above equations for $T = 0$ gives a stable region for the solutions corresponding to bump formations for different values of the load α, the sparsity a and the retrieval asymmetry parameter R, shown in Fig.1. As can be seen, the sparsity of the code a enhances the SAS effect, although it is also observed for $a = 0$. As we expected, the asymmetry factor R between the stored and retrieved patters is very important in order to have spatial asymmetry. The diagram in Fig.1 shows a clear phase transition with R. Only for intermediate values of R, the bump solution exists.

The dependence of the critical storage capacity α_c on the asymmetry parameter R is shown in Fig.2. The figure presents a drastic drop of α_c at the transition from homogeneous retrieval (symmetric) state to spatially localized (bump state). Effectively only the fraction of the network in the bump can be excited and the storage capacity drops proportionally to the size of the bump.

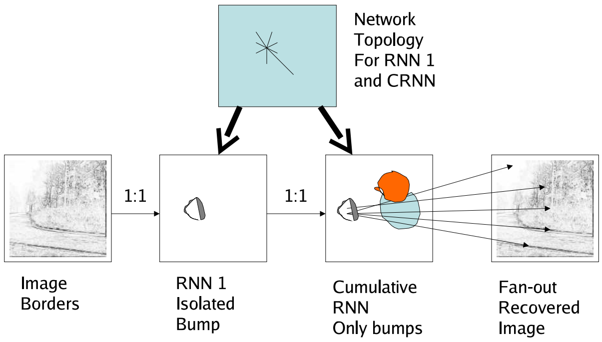

Fig. 3. Schematics of the two stage pattern storage. The first stage forms a bumps in RNN 1 that contains few patterns. The second stage records this bump in a cumulative RNN with high capacity. The fan-out recovers the original image.

3 Computer experiments

As we mentioned in the Introduction, we have performed computer experiments to show that the capacity of the network can be increased to the limits typical for sparsely connected network.

As a test we used the natural image database of van Hateren [11], as well as the Waterloo set of images. In order to make the simulation easier, we decreased the resolution of the images down to 256 by 256 points. We have used randomly selected subsets of up to 300-700 images of van Haterens database and also consecutive images of that database. No difference in the results, concerning the capacity of the network has been found. Waterloo dataset is very small and was used as a reference set together with the van Heteren database. Once again it was found that the results do not depend of the database source.

For the de-correlation of the image, we used a border detector, converting the original image into an image that represents its borders. The simulations show that the exact type of the border detector does not affect the results. Each pixel of the resulting image is represented by a neuron in the network. We tested binarised version of the images borders as well as the discrete (gray) values of the pixels in order to train the network. Both of them give approximately the same results.

We applied a two stage learning procedure Fig.3. In the first stage we trained a Hebbian network with spatially dependent connectivity and probability of connections $p(r) \propto const + e^{-r^2/2\sigma^2}$, where r is the radius of connectivity. The trained network contains few patterns, or even only the pattern in question. Then

Fig. 4. The contours of Lena and the bump formation with sparsity $a = 0.005$ and $r = 15$. The bump represents a compact area.

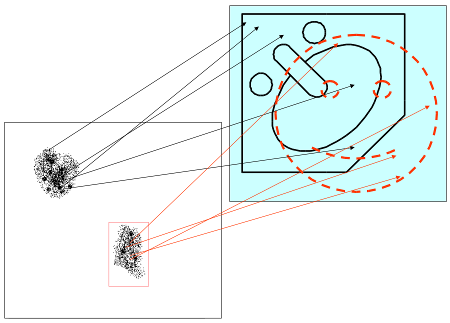

Fig. 5. Fanout to recover the image from its bump. Usually the bump is unique, although up to 4 separate bumps are easily observable.

we raised the threshold in order to observe a bump with determined sparsity. The bump formation is sufficiently noise resistant and stable in order to reconstruct the complete image. An example of such a bump is given in Fig. 4. The location of the bump depends exclusively on the topology of the network and on the pattern. For a typical image, usually only one of few (two to three) bumps are formed.

In the second stage, we used the bump in order to train a Hebbian network with exactly the same topology, that contains only bumps from a variety of the images. We referred this network as a Cumulative Recurrent Neural Network (CRNN).

In a parallel work [12] we have shown that the image can be restored effectively from its bump. This gives us an opportunity to save only the bump in a network and to associate another feed-forward network in order to restore the image in question.

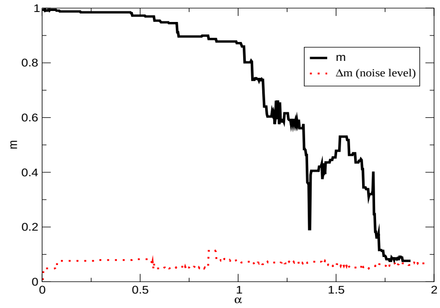

Fig. 6. Typical behavior of the order parameter m, normalized to the sparsity of the pattern as a function of the network load α. The dashed curve is the "recovering" of an arbitrary not memorized pattern, e.g. the noise level.

Roughly speaking the fan-out part of the network connects a dilation of the bump/bumps to all the neurons of the fan-out network. The rest of the neurons, i.e. the neurons outside of the bump are not connected to the fan-out layer. Thus if the areas of two different bumps do not intersect, the recovered patterns clearly do no interact, Fig.5.

The intersecting bumps, when they are orthogonal enough, do not cause a problem. In other words, with capacity up to the critical capacity α_c, we have a good recovery of the initial pattern. The details of this results are presented in Ref.[12]. The noise level is usually less than the dashed line in Fig.6 that is a small quantity compared to the signal.

In order to keep the network in regime $C \propto N$, the connectivity of the system must be large enough. In these simulations we used $C = 300$, that is suitable for computer simulations. It fits well into the computer memory, it is large enough in order to be in the range $\propto N$ and shows a rich behavior. For this value of C, the correlation radius r can vary from 10 to the size of the image.

For large values of r, no bump formations can be observed. Therefore, one can expect that the capacity of the network will reach its asymptotic limit proportional to $1/a|\log a|$. However, the pattern is very unstable and susceptible to noise, which makes the capacity close to zero, Fig.7(right).

When r is very small, the network effectively converts into locally connected network. In this case the capacity is small, Fig. 7(left), but not too small because the pattern are spatially correlated (the borders are continuous curves) and therefore the bump is well formed. Thus, different images are kept in different parts of the cumulative CRNN and they are essentially orthogonal.

For intermediate levels of the correlation radius r, one can achieve capacity close to the theoretical capacity $\alpha = 1.58$ for sparse random network, Fig. 6.

The critical capacity versus the load is shown in Fig.8. One can see that the critical capacity has a maximum near $r = 20$ and in general it is very high. It is of the same order as the critical capacity of a very sparse network, up to the

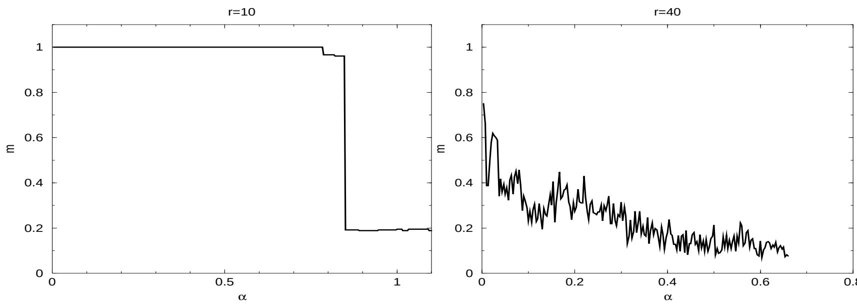

Fig. 7. The bump capacity with too small (left) and too large (right) correlation radius r.

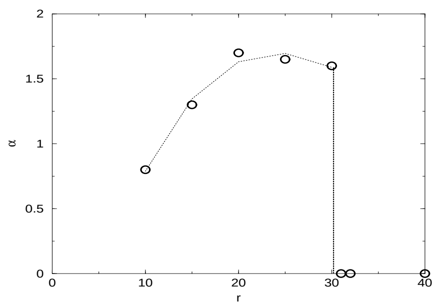

Fig. 8. Non-trivial behavior of the critical storage capacity during the bump formation phase. For $r = 31.3, \alpha_c$ drops to 0.

correlation radius where the bumps disappear and the capacity drops sharply to zero.

4 Conclusion

In this paper we have shown that by using a special two-stage learning procedure and a bumpy network, one can increase the storage capacity of this network even with Hebbian learning rule. This feature is based on the theoretical background, showing that the location of the bump is only slightly dependent on the patterns already stored in the network. Therefore, one can use the bump as an intrinsic characteristic of the image and the topology of the network . The architecture is promising when another, fan-out layer is added to it, because this allows a total recovering of the image with relatively few connections.

Acknowledgments

The authors acknowledge the financial support from the Spanish Grants DGI.M. CyT. FIS2005-1729, Plan de Promoción de la Investigación UNED and TIN 2004–07676-G01-01. We also thank David Dominguez for the fruitful discussion of the manuscript.

References

1. Y.Roudi and A.Treves, JSTAT, P07010, 1 (2004).
2. Y.Roudi and A.Treves, cond-mat/0505349.
3. K.Koroutchev and E.Korutcheva, Preprint ICTP, Trieste, Italy, IC/2004/91, 1 (2004);
   K.Koroutchev and E.Korutcheva, Central Europ. J.Phys., **3**, 409 (2005);
   K.Koroutchev and E.Korutcheva, Phys.Rev.E **73** (2006)No 2.
4. A.Anishchenko, E.Bienenstock and A.Treves, q-bio.NC/0502003.
5. V.Breitenberg and A.Schulz, *Anatomy of the Cortex*, Springer, Berlin, 1991.
6. D.J.Watts and S.H.Strogatz, Nature, **393**, 440 (1998).
7. D.J.Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness(Princeton Review in Complexity)*(Princeton University Press, Princeton, New Jersey, 1999).
8. J.Hopfield, Proc.Natl.Acad.Sci.USA, **79**, 2554 (1982).
9. M.Tsodyks and M.Feigel'man, Europhys.Lett., **6**, 101 (1988).
   Stat.Phys., **14**, 565 (1994).
10. D.Amit, H.Gutfreund and H.Sompolinsky, Ann.Phys., vol.173,(1987),pp.30-67.
11. van Hateren J. H. and van der Schaaf A, *Independent component filters of natural images compared with simple cells in primary visual cortex.* in Proc.R.Soc.Lond. B, 265:359-366, 1998.
12. K.Koroutchev and E.Korutcheva, in the Proceedings of the 9th Granada Seminar on Computational and Statistical Physics, AIP, 2006.