

# Learning Block Memories with Metric Networks

Mario Gonzalez, David Dominguez, and Francisco B. Rodriguez

**Abstract**—An attractor neural network on the small-world topology is studied. A learning pattern is presented to the network, then a stimulus carrying local information is applied to the neurons and the retrieval of block-like structure is investigated. A synaptic noise decreases the memory capability. The change of stability from local to global attractors is shown to depend on the long-range character of the network connectivity.

**Keywords**—Hebbian learning, image recognition, small world, spatial information.

## I. INTRODUCTION

WHILE it is known that the most efficient arrangement for storage and retrieval of patterns by an attractor neural network as a whole (global information) is the random network[1], connections in real brains do not appear to be fully random. The majority of connections appear to occur between nearby regions, suggesting a small-world topology[2],[3].

In this work a model of sparsely connected Hopfield-type neural networks [4],[5] on the small-world topology of Watts-Strogatz is presented[6],[7]. It is studied in deeper detail the advantages of the small-world network topology going from the one dimensional ring to the random graph[8],[9], with emphasis on networks capable of recovering global information where local stimulus is relevant or only block information is at disposal for the network. This can be successfully applied in pattern recognition, such as image recognition, or any type of pattern carrying local spatial information[10].

The response of a network to a given input stimulus leads to a particular configuration of the neural activity. Local order can emerge if the stimulus has some neighbourhood structure, and the topology of the network favors stronger connectivity with nearest neurons than between neurons in far regions. The block-like attractor is explored to study in which conditions (parameters), its dynamical behaviour favors retrieval of local spatial information.

## II. THE MODEL

In order to describe the neural activity the following model is used. A pattern is presented to the network and the neurons  $\sigma_i$  are initialized in one of two states,  $\pm 1$  (active/inactive). The state of a neuron  $\sigma_i$  is updated in time  $t$  through the following equation:

$$\sigma_i^t = \text{sign} \left( \sum_i h_i^{t-1} - \theta \right), \quad h_i^t \equiv \sum_j J_{ij} \sigma_j^t \quad (1)$$

Mario Gonzalez, mar.gonzalez03@estudiante.uam.es

David Dominguez, david.dominguez@uam.es

Francisco B. Rodriguez, f.rodriguez@uam.es

All of the authors are members of the GNB at EPS, Universidad Autonoma de Madrid, 28049 Madrid, Spain.

where  $h_i$  is the postsynaptic field arriving at neuron  $\sigma_i$ . Here the sign function is defined as:  $\text{sign}(z) = 1$  if  $z \geq 0$ , and  $\text{sign}(z) = -1$  if  $z < 0$ . where  $K_i$  is the number of connections of each neuron and  $W_{ij}$  represents the weights of the synapses between neurons  $i$  and  $j$ . The variable  $\theta$  is the firing threshold which is considered to be zero. We use synchronous update and asymmetric weights. The synaptic couplings between neurons  $i, j$  are  $J_{ij} \equiv C_{ij} W_{ij}$ , where  $\mathbf{C} = \{C_{ij} \in \{0, 1\}\}$  is the topology matrix and  $\mathbf{W} = \{W_{ij}\}$  are the learning weights.

The topology matrix splits in local and random links. The local links connect each neuron to its  $K_L$  nearest neighbours, in a closed ring. The random links connect each neuron to  $K_R$  others uniformly distributed along the network. Hence, the network degree is  $K = K_L + K_R$ . The network topology is then characterized by two parameters: the *connectivity* ratio, and the *randomness* ratio, defined respectively by:

$$\gamma = K/N, \quad \omega = K_R/K, \quad (2)$$

where  $\omega$  plays the role of a rewiring probability in the *small-world* model[2].

The synaptic weights  $W_{ij}$  are given by

$$W_{ij} = cW_{ij}^r + (1 - c)\xi_i\xi_j \quad (3)$$

The term  $W_{ij}^r$  is randomly generated to be either  $+1$  or  $-1$  with equal probability, multiplied by a factor  $c \in (0, 1)$  which is the load rate, which accounts for all the previous synaptic processing, including both short-term and long-term memory of the network. The second term describes the Hebbian learning of a given pattern multiplied by  $1 - c$ . For a given  $c$ , there is a competition between the learning pattern  $\vec{\xi} \equiv \{\xi_i\}_i^N$  and the noise played by the random term. The pattern the network is learning is the picture of Lena in Figure 1 - left.

We use a mesoscopic variable  $m_l^t$  to describe the overlap between the neural activity of block  $l$  of information, and the pattern activity in this same piece of information, at time  $t$ . The *block - overlap* restricted to the block  $l$  is:

$$m_l^t \equiv \frac{1}{N_l} \sum_{i \in l} \xi_i \sigma_i^t, \quad (4)$$

at time step  $t$ . We can define averages over blocks as:  $\langle f_l \rangle_b \equiv \frac{1}{b} \sum_{l=1}^b f_l$ .

The relevant order parameter are the mean *overlap* between the neural states and the pattern,  $m^t$ , and the blocks-variance  $v^t$ , given by (dropping the time index  $t$ )

$$m \equiv \langle m_l \rangle_b, \quad v \equiv \langle m_l^2 \rangle_b - m^2. \quad (5)$$

Note that  $m$  is the usual *global* overlap, also written as  $m \equiv \frac{1}{N} \sum_i \xi_i \sigma_i$ . The standard deviation, one names *block* overlap is  $\delta = \sqrt{v}$ . If the size of the blocks are taken  $L = 1$ , then

the  $b = N$  blocks, with  $m_l = \pm 1$ , carry no macroscopic order. On the other hand, if there is only one block  $b = 1$ , then  $v = 0$ . However, if the size is large but  $1 \ll L \ll N$ , the blocks carry local information. We also define  $m$  as the similarity between the learned and recovered pattern, and  $\delta$  as their local similarity.

### III. RESULTS OF SIMULATION

Figure 1 - left shows the original pattern learned by the network which is the picture of Lena Söderberg, a standard test image used in digital image processing. The picture is  $256 \times 256$  pixels, it has been binarized and properly formatted to be either  $\pm 1$  by pixel. We want to study the stability and the attractor properties of a block state of the network, characterized by a spatial partition of the pattern in a correct zone and an inverted zone. The correct zone has the neuron states identical to the learned pattern, while the inverted zone has the neurons states in opposition to the pattern: each neuron positioned at the same site as the pattern is switched off, i.e. the active neurons become inactive and other way round. It is represented in Figure 1 - center.

This state has vanishing overlap with the original pattern, however one sees that it still carry information about Lena! Most of the work found in the literature about memory networks with metrics [1],[6],[7],[8] are focused only on the global retrieval of a pattern, does not considering the possibility of spatially correlated states. To study the stability of the block states, one should start the network evolution precisely with the block, so that  $m^{t=0} = 0$  but  $\delta^{t=0} = 1$ . Then, if the network stays at this block state, or if it goes close to it, it means that there is a block phase which is stable. However, this could be a marginal stability, with a very narrow attractor basin. To check for the size of the attractor basin, one must verify different initial conditions. For instance, the initial configuration considered here in this paper has  $m = 0$ ,  $\delta = 0.2$ , which is a strongly noisy 2-blocks structure of overlaps, as can be seen in Figure 1 - right.

The network ( $N = 65,536$ ,  $K = 64$  in average) starts in a noisy 2-blocks Lena, and the neural dynamics leads to a stationary state. Two cases are plotted in the Figure 2: at left, the network has  $\omega = 0.1$ ,  $c = 0.74$  at right it is  $\omega = 0.2$ ,  $c = 0.8$ . One sees that the attractor for a more local topology (small  $\omega$ ) and with moderate loading (not so high  $c$  rate) is the block state (Fig.2 - left). For larger number of long-range connections and higher loading, the attractor is the global state (or the negative of it, as in Fig.2 - right). The evolution in time for the network with block attractor is depicted in Figure 3. One sees that the blocks are fulfilled in the first  $t = 10$  time steps, and it stays forever in the B phase, except for small random fluctuations. The evolution for the network with the global attractor is in Figure 4. One sees that the network quickly evolves into a global phase, with a completion of 93 percent of the learned pattern,  $m \approx 0.93$ , according to the Figure 2 - right.

In Figure 5 a phase diagram describes the regions of block (B), global (G) and stationary (Z) states. One can conclude from this figure that block activity appears for values of  $c$



Fig. 1. Left: Lena - original. Center: Lena 2-blocks. Right: Lena 2-blocks,  $v = 0.2$ .



Fig. 2. Recovered patterns. Left:  $\omega = 0.1$ ,  $c = 0.74$ , right:  $\omega = 0.2$ ,  $c = 0.8$ .

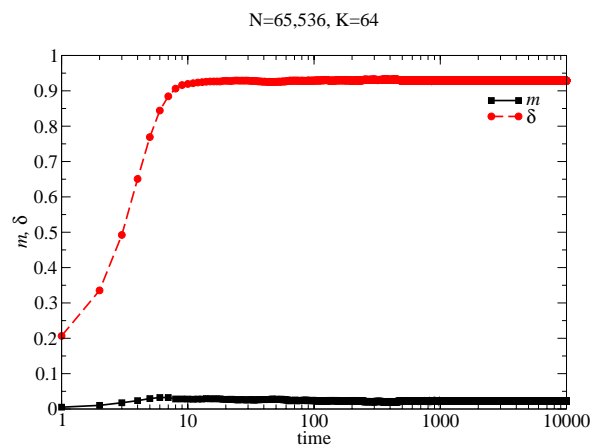


Fig. 3. Evolution of the network into a block phase:  $\omega = 0.1$ ,  $c = 0.74$ .

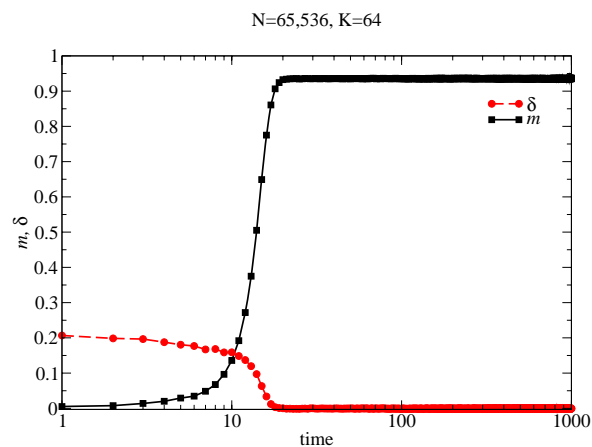


Fig. 4. Evolution of the network into a global phase:  $\omega = 0.2$ ,  $c = 0.8$ .

less than approximately 0.84, and for  $\omega$  not greater than 0.3. One can observe a phase transition between global and block retrieval.

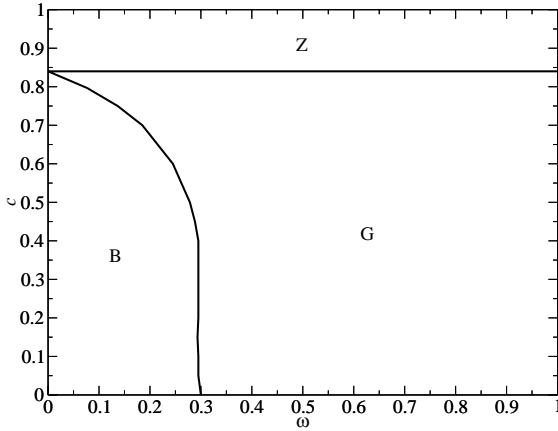


Fig. 5. Phase Diagram for block activity. B: block activity region. G: global activity region, Z: stationary region.

#### IV. SIGNAL TO NOISE THEORY

The simulation results presented in the previous section can be supported by a straightforward theory. The theory discussed here is based in a signal to noise ratio approximation. Let the neurons be distributed within blocks  $l$ , successively with positive and negative activities,  $m_l = m_{\pm}$ . Then, following Equations (5), the block activities can be written as

$$m_l = m + y_l \delta, \quad (6)$$

where  $y_l \doteq \pm 1$  (according to the block) is a random variable.

The local field, Equation (1), with the Equation (3) for the synapses, can be separated in a signal and a noise terms,

$$h_i \equiv (1 - c)K\xi_i m_i + c\Omega_i \quad (7)$$

where  $m_i \equiv \frac{1}{K} \sum_{j \in \{i\}} \xi_j \sigma_j$ ,  $\Omega_i \equiv \sum_{j \in \{i\}} W_{ij}^r \sigma_j$  are the activity restricted to the neighbours  $\{i\}$ , and the synaptic noise, respectively.

There are local and random neighbours for each neuron, hence the signal term itself splits in localized and randomized terms, namely

$$m_i = \frac{K_L}{K} m_i^L + \frac{K_R}{K} m_i^R, \quad (8)$$

with  $m_i^{L,R} \equiv \frac{1}{K_{L,R}} \sum_{j \in K_{L,R}} \xi_j \sigma_j$  where  $L$  and  $R$  are the local and random sets of neighbours, respectively, of the neuron  $\sigma_i$ .

From Equation (4), whenever the neighbours belong to a block, the localized field depends on its block overlap,  $m_l$ . On the other hand, the randomized field is a sample of a global field, which does not depend on the block. Using the definition in Equation (2), one arrive to an approximation for the local field of neurons in the block  $l$ ,

$$\xi h_l \equiv (1 - c)K[\omega m + (1 - \omega)(m + y_l \delta)(1 - \gamma b)] + c\Omega \quad (9)$$

where the correction term  $(1 - \gamma b)$  accounts for the boundary effects between  $m_{\pm}$  blocks.

The equation for the the block-activity is then  $m_l = \langle \text{sign}(\xi h) \rangle_{\Omega}$ , where the average in the angular brackets are over the noise  $\Omega$ . But from the Equation (6), after averaging over the  $y_l$  one gets

$$\begin{aligned} m &= \langle m_l \rangle_y = \langle \text{sign}(h) \rangle_{y,\Omega} \\ \delta &= \langle y m_l \rangle_y = \langle y \text{sign}(\xi h) \rangle_{y,\Omega}, \end{aligned} \quad (10)$$

The average over  $\Omega$  stands for the noise distribution.

This noise is Gaussian distributed,  $\Omega \doteq N(0, \Delta^2)$  [5]. Its variance is given by the sum of random and local terms,  $\Delta^2 = \text{Var}(\Omega_i) = \omega \Delta_r^2 + (1 - \omega) \Delta_l^2$ . Neglecting the *feedback* terms, it is  $\Delta^2 = K$ . This approximation is valid in the limit of strongly diluted networks ( $K \ll N$ ) [10]. However, for local connections, even extreme dilution do not eliminate the feedback, and  $\Delta$  needs more precise calculations, which is outside the scope of the present work.

The continuous transition from the G to the Z phase may be analysed by taking first  $\delta = 0$  in the Equations (10), which gives  $m = \langle \text{sign}[(1 - c)K m + c\Omega] \rangle$ , then expanding around  $m \sim 0$ . It gives the constant line:  $c = [1 + \sqrt{\frac{\pi}{2K}}]^{-1}$ , which coincides with border G-Z plotted in Figure 5. The transition between B and G phases is not continuous, so no expansion is possibly, but the equation:  $\delta = \langle \text{sign}[(1 - c)K(1 - \omega)\delta + \Omega] \rangle$ , is similar to the previous equation for  $m$  except that it depends on  $\omega$ . The finite solution  $\delta > 0$  is stable only if  $c = [(1 - \omega) + \sqrt{\frac{\pi}{2K}}]^{-1}$ , which fits well also with the phase diagram in Figure 5.

#### V. CONCLUSION

An attractor neural network with a metrical topology, (i.e. a small-world with local and random connections) was studied in this paper. The overlap between a learning pattern and the neuron states is the usual global parameter which measures the network retrieval ability. When the connections are preferentially local, however, there are spatial correlations between neurons which allow for states retrieving blocks of a pattern. It is observed here that these block states are stable, with a large basin of attraction.

Simulations for the dynamics of the global and block overlaps are presented, and a diagram for the transitions between global, block and zero-retrieval phases is built. A theoretical set of macroscopic equations is obtained, and their fixed point solutions agree quite well with the simulation results.

The present proposal of a block-like structure of patterns, can be extrapolated to other types of topology. It could be closely related to biological brain systems, on the one hand, where different sensory blocks of patterns (arising from several cortical structures) may be independently retrieved. On the other hand blocks may represent incomplete pieces of information which can be used to codify images (as shown here) or other kind of signals (voice, fingerprints, genetic code, etc). These are subjects of future research.

#### ACKNOWLEDGMENT

This work was supported by TIN 2004-04363-CO03-03, TIN 2007-65989 and CAM S-SEM-0255-2006.

## REFERENCES

- [1] Patrick N. McGraw and Michael Menzinger, *Topology and computational performance of attractor neural networks*, Phys. Rev. E 68, 047102, 2003.
- [2] Watts, D. J. and Strogatz, S. H., *Collective dynamics of 'small-world' networks*, Nature, 393-6684, pages 440-442, 1998.
- [3] R. Albert and A. Barabasi, *Statistical mechanics of complex networks*, Rev. Mod. Phys., 74-1, pages 47-97, 2002.
- [4] D. R. Paula, A. D. Araujo, J. S. Andrade Jr., H.J. Herrmann and J. A. C. Gallas, *Periodic neural activity induced by network complexity*, Phys. Rev. E 74, 017102, 2006.
- [5] D. J. Amit, *Modelling Brain Function: The World of Attractor Neural Networks*, Cambridge University Press, 1989.
- [6] G. Morelli, G. Abramson and M. N. Kuperman, *Associative memory on a small-world neural network*, The European Physical Journal B 38-3, pages 495-500, 2004.
- [7] T. Nikolettopoulos, A. C. C. Coolen, I. Perez Castillo, N. S. Skantzos, J. P. L. Hatchett and B. Wemmenhove, *Replicated transfer matrix analysis of Ising spin models on small world lattices*, J. Phys. A: Math. Gen. 37, pages 6455-64754, 2004.
- [8] J. P. L. Hatchett, N. S. Skantzos, and T. Nikolettopoulos, *Dynamic rewiring in small world networks*, Phys. Rev. E 72, pages 066105, 2005.
- [9] D. Dominguez, K. Koroutchev, E. Serrano and F. Rodriguez, *Associative memory: information and topology*, Research in Computing Science, Mexico, pages 39-48, 2006.
- [10] D. Dominguez, *Block information and topology in memory networks*, CP887, Coop. Behavior in Neural Systems, Granada, pages 107-114, 2007.



**Francisco B. Rodriguez.** received his degree in Applied Physics in Universidad Autonoma de Madrid. He then worked at the Nijmegen University in Holland and the Institute for Nonlinear Science, University of California, San Diego. Since 2002 he is profesor titular at the EPS, UAM.



**Mario Gonzalez.** Postgraduate student at Escuela Politecnica Superior (EPS), Universidad Autonoma de Madrid (UAM), Spain. His researches include neural networks and information theory. Member of GNB-UAM (Grupo de Neurocomputacion Biologica).



**David Dominguez** has a PhD in physics from the UFRGS in 1993. He worked at the PUC/RS and FAPA (Porto Alegre), at the KUL (Leuven), at the INTA, URJC and UCM (Madrid). Since 2001 he is a professor at the EPS-UAM.