A scale free neural network for modelling neurogenesis

Juan I. Perotti, Francisco A. Tamarit [1] and Sergio A. Cannas [1]

Facultad de Matemática, Astronomía y Física, Universidad Nacional de Córdoba
Ciudad Universitaria, 5000 Córdoba, Argentina

Abstract

In this work we introduce a neural network model for associative memory based on a diluted Hopfield model, which *grows* trough a neurogenesis algorithm that guarantees that the final network is a small–word and scale–free one. We also analyze the storage capacity of the network and prove that its performance is larger than that measured in a randomly dilute network with the same connectivity.

Key words: Neural networks, Scale–free networks, Strong dilution
PACS: 84.35.+i

Artificial neural networks have been a subject of intense study during the last decades, mainly because their applications to both mathematical modelling of biological nervous systems and artificial intelligence. A cornerstone in the development of artificial neural networks was the work by Hopfield [1] where he showed, through a very simplified model, how distributed associative memory can emerge as the result of *a)* a non linear universal neural dynamics and *b)* a plastic network of synaptic interactions. Although highly unrealistic from the biological point of view, the *Hopfield model* was able to emulate some associative processes used by animals when storing and retrieving information. Moreover, it linked this field and the statistical mechanics of disordered magnetic systems.

Since then, several modifications of the Hopfield model have been proposed in order to endow it with some biological features, and many of these modifications concerned the structure of the underlying connectivity network. The

Email addresses: juanpool@gmail.com (Juan I. Perotti,), tamarit@famaf.unc.edu.ar (Francisco A. Tamarit), cannas@famaf.unc.edu.ar (and Sergio A. Cannas).
[1] Member of Consejo Nacional de Ciencia y Técnica, Argentina

Hopfield model assumed that all the neurons are fully and symmetrically connected, two biologically undesired features that make the model analytically tractable. The memory load of the system, defined as the ratio between the number of stored bits and the total number of synapses in the network $\alpha = \frac{pN}{N(N-1)} \sim \frac{p}{N}$ presents, for uncorrelated patterns, a critical value $\alpha_c \approx 0.138$, above which the network cannot longer recover the stored patterns. Derrida, Gardner and Zippelius [2] presented an ultra–diluted Hopfield model (hereafter referred as DGZ network), where every node is connected asymmetrically to just C other random chosen nodes, with $C \ll \ln N$. In this case, the storage parameter α takes the form $\alpha = pN/CN = p/C$ and its critical value increases to $2/\pi$. In other words, not only associative memory appears as a robust property against dilution and asymmetry, but those properties improve it. Later on, Arenzon and Lemke [3] showed, through very extensive numerical simulations, that this improvement actually holds for the more realistic and feasible case $C \ll N$.

Montemurro and Tamarit [4] presented a modification of the ultra–diluted model (hereafter referred as MT network) in which the synapses of an initially fully connected network are removed systematically through an extremal pruning strategy. For every node of the network, all but the C strongest synapses are removed, where strongest refers to the absolute value of the synaptic weights. They showed that, even in the less restrictive limit $C \ll N$, α_c diverges logarithmically with N and therefore the storage capacity of the network never disappears in the limit $N \to \infty$. This shows that the storage capacity of a neural network can be highly improved by correlating the underlying connectivity of the nodes to the synaptic weights, or equivalently, to the activity of the network under working conditions.

Both the above mentioned models (DGZ and MT) leads to an underlying connectivity network which is a fully random graph. However, it has been shown in the last years that networks of interacting entities appearing in biological systems are not random, but display complex properties, such us small–world character [5] and scale–free distribution of connectivity degree [6]. Moreover, in a recent study on the human brain activity using fMRI, Eguiluz and coworkers [7] showed that a functional network, defined through the correlations between the activities of different areas, displays small world and scale free properties. Although such network is based on a heavily coarsening process of the information in the brain, a recent work by Kim [8] support the possibility that such properties of the functional network reflect a similar structure of the underlying neural networks. Clearly, such structure is not expected to be a fully random graph, since it is highly determined by some learning process. In other words, the wiring and rewiring of a brain is known to be highly correlated with the activity of the neurons during some stage of the learning process. Moreover, it has been recently realized that adult neurogenesis, i.e., the incorporation of new neurons into already developed neural structures, is

another form of adaptation to environmental changes, i.e., learning [9]. Hence, it is of interest to analyze how the performance of basic neural networks models is affected by the incorporation of learning rules which work during the growing stage of a network that develops a complex structure.

Stauffer and coworkers [10,11] analyzed recently the storage capacity of a Hopfield model defined on a scale free network constructed following the Barabási–Albert algorithm [6]. Although this network reflects the observed structure of nervous systems, it is important to stress that its development does not take into account the activity of the neurons. In other words, it presumes a complete genetic imprinting of the topology of the neural network, irrespectively of any pattern of brain activity. Despite this unrealistic feature, the model, as occurs with the MT model, predicts an unlimited storage capacity.

In this work we introduce a neural growing algorithm in which new neurons link others taking into account the pattern of activity of the network. As will become clear soon, this algorithm produces small–world scale–free networks characterized by a large cluster coefficient, a small average path length and a power–law connectivity distribution, reminiscent of those neural nets observed in some regions of large nervous systems.

Our model is a modified version of the Hopfield one where each neuron is modelled by an Ising variable $S_i = \pm 1$ (representing the firing $(+1)$ and resting (-1) states) and the synaptic architecture is constructed following a Neuro–Genesis Algorithm (NGA). We start with a small network of size N_1 constructed according to the MT algorithm, i.e., optimizing the pruning process in such a way that only the C strongest synapses survive after dilution. This deterministic process can be thought as mimicking the genetic initial imprinting of every natural nervous system. The p stored patterns are denoted as $\{\xi_i^\mu\}$ with $i = 1, \ldots, N_1$ and $\mu = 1, \ldots, p$, and the surviving synapses are constructed following the Hopfield rule:

$$w_{ij} = \sum_\mu \xi_i^\mu \xi_j^\mu \quad \text{for} \quad i \neq j \qquad \text{and} \quad w_{ii} = 0 \tag{1}$$

The quantities ξ_i^μ are independent random variables taking the values ± 1 with equal probability. Once the initial net has been constructed, we add sequentially new neurons and connect each of them to C of the previously existing neurons, in such a way that the probability of linking the new neuron to a preexisting neuron i is given by

$$\Pi_i(t) = \frac{\sum_j |w_{ij}| + |w_{ji}|}{\sum_{jk} |w_{jk}| + |w_{kj}|} \tag{2}$$

In other words, we consider the numerator $\sum_j |w_{ij}| + |w_{ji}|$ as a measure of the

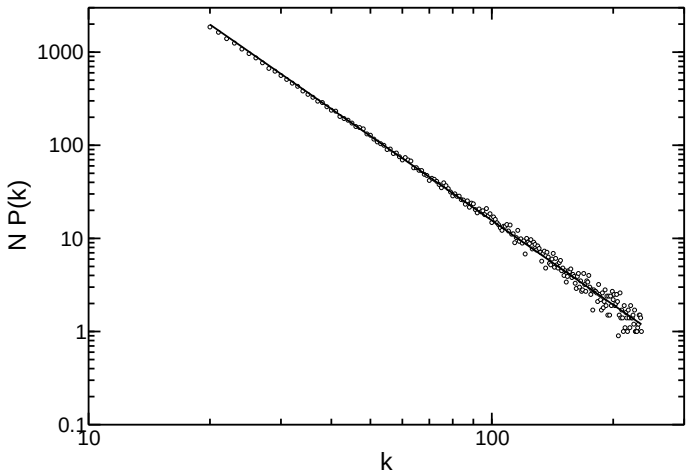

Fig. 1. Connectivity distribution function for $N_1 = 60$, $N_2 = 20000$ and $p = 20$.

activity of neuron i, irrespectively of the inhibitory or excitatory character of the synapses. Note that the sum runs only over the connected neurons, or in terms of graphs, over the nearest neighbors. After creating a new neuron j we provide it with p new random binary variables ξ_j^μ ($\mu = 1, \ldots, p$) that represent the states of this neuron on each stored pattern. Once the neighbors are chosen the new synapsis are constructed according to the Hebbian rule (1). Note that the synapses are asymmetric and more than this, they are always ingoing connections. The biological appealing of this algorithm is straightforward. In what follows we will analyze both the topology of the emerging network and its recognition capacity in order to compare it with the already presented modifications of the Hopfield model.

In Fig. 1 we display, in a double logarithmic plot, the connectivity distribution function $NP(k)$ of a neural network of $N_2 = 20000$ neurons when the system is initialized with $N_1 = 60$ neurons and $p = 20$ configurations are stored. The power law behavior observed is a clear signature of the scale–free character of the net, and the best fitting (continuous line) corresponds to an exponent $\gamma = 2.98 \pm 0.02$, which is numerically indistinguishable from the exponent obtained with the Barabási–Albert algorithm [6]. It is important here to stress that these results are almost independent of the value of p but as N_1 increases the topology gains a more complicated structure (to be published elsewhere).

In order to characterize the *small word* character of the net, we have also analyzed both the clustering coefficient c and the mean free path l. The clustering coefficient of a given neuron i measures the connectivity of its nearest–neighbors and it is defined as the probability of any two direct neighbors to be connected,

$$c_i = \frac{E_i}{k_i(k_i - 1)}\,, \tag{3}$$

where k_i is the connectivity of neuron i and E_i is the total number of existing synapsis between its direct neighbors. The global clustering coefficient c is then the average of all the clustering coefficients of the network. Topologically, a high value of c indicates a tendency to form triangles (note that in our case, c does not take into account the weight of the synapsis). We have measured the mean cluster coefficient of the final neural network starting with different initial sizes and different numbers of stored memories. In all these cases we have verified that the cluster coefficient is close to $c = 5.2 \times 10^{-2}$, which is almost twice the cluster coefficient corresponding to a non–directed Barabási–Albert network.

The *mean path length* l between two nodes, defined as the number of edges along the shortest path connecting them, is another relevant topological property of a graph, and in particular for scale–free nets the value of l is relatively small, growing logarithmically as their sizes increases. We have calculated l for different values of the parameters, finding in all the cases that is of the order or smaller than six, and that it decreases as the connectivity of the network C increases. The large clustering coefficient c together with the small average path length l confirm the small world character of the network.

Next we present a numerical study of the storage capacity of the neural network created by the NGA. To do that we follow the usual protocol: after constructing a net of N_2 neurons with p stored patterns, we initialize the system in one these memories, let's say the ν-th. Next, we let the system evolve using a sequential deterministic dynamics governed by the rule,

$$S_i(t + \Delta t) = \mathrm{sign}(h_i(t)) \qquad \text{with} \qquad h_i(t) = \sum_{j \neq i}^{N_2} w_{ij} \, . \tag{4}$$

until it reaches a stationary regime and from then on we measure the temporal average m of the overlap of the system with the pattern ν, defined as,

$$m_\nu(t) = \frac{1}{N} \sum_{i=1}^{N} s(t) \xi_i^\nu \, . \tag{5}$$

The closer m approaches to one (zero), the more stable (unstable) is the corresponding stored pattern.

We first compare the performance of the NGA network with those obtained with the optimization pruning process (MT) and with the randomly ultra diluted (DGZ) networks. In Fig. 2 we plot m *vs.* $\alpha = p/\langle C \rangle$ for networks of $N_2 = 20000$ neurons and fixed connectivity $C = 20$. Note that the curve corresponding to the NGA has an initial rapid decay, indicating a poorer behavior than that observed for the MT networks but, what it is important

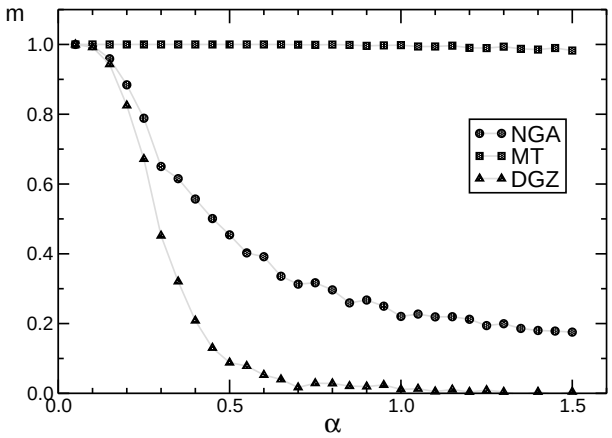

Fig. 2. Comparison between different network topologies of the behavior of m vs α: MT (full circles), NGA (full squares) and DGZ (full triangles). In all the cases the system size and the connectivity are $N_2 = 10000$ and $C = 20$ respectively. For the NGA case the initial network has $N_1 = 50$ neurons.

here to highlight is the fact that it largely improves the performance of the random DGZ network. Furthermore, the decay of $m(\alpha)$ seems to be slow enough to guaranty that the model can retrieve the storage patterns for any finite value of α, though with a poor load parameter m. The performance of the NGA is also poorer than that observed when, after creating a Barabási–Albert scale–free network a Hopfield matrix is imprinted on it [10], but this biologically unrealistic process implies the existence of symmetric couplings, a property that it is known to improve the storage capacity. When compared with the asymmetric version of this last model the NGA is slightly better [11].

In Fig. 3 we plot m vs α for $N = 10000$, $C = 20$ and different sizes of the initial network (increasing from bottom to top). As can be seen, the larger the initial network, the better the retrieval performance of the network, a fact that reveals the increasing role that play the optimization pruning process on the overall behavior of the system. This is an biologically appealing result that somehow explain the tendency observed in nature to create complexity along specialization by enlarging the size of the embrional brain, preserving almost the unchanged the connectivity and reducing the role of neurogenesis on both adaptation and learning [12].

Summarizing, we have introduced a growing neural algorithm that creates neurons taking into account the activity pattern of the previously existing network. Starting with a very small optimized net that emulates the basic genetic structure of any nervous system (which in fact has been submitted to the optimization process of natural selection) the network is able to recover information even for large load parameter. Furthermore, the load performance deteriorate slow enough to assure at leat an unusual large storage capacity.

This work was partially supported by grants from CONICET (Argentina),

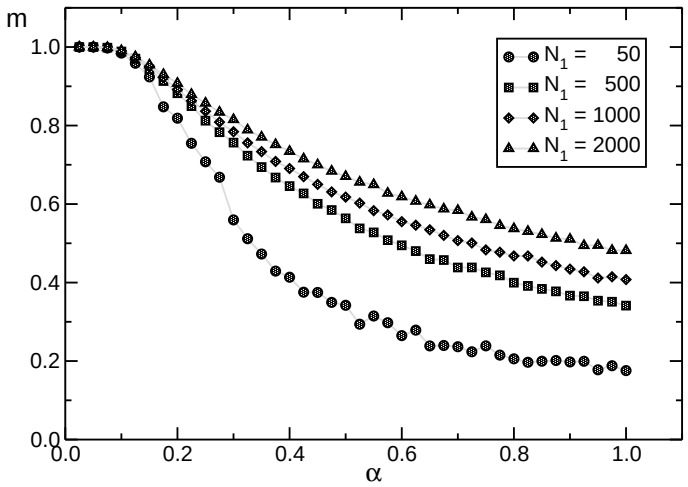

Fig. 3. m *vs.* α for a NGA network of $N_2 = 20000$ neurons with $\langle C \rangle = 20$ and different sizes of the initial MT network N_1: 50 (full circles), 500 (full squares), 1000 (full diamonds) and 2000 (full triangles).

Agencia Córdoba Ciencia (Argentina) and SECyT (Universidad Nacional de Córdoba, Argentina).

References

[1] J. J. Hopfield, Proc. Nat. Acad. Sci. USA **79**, 2554 (1982).

[2] B. Derrida, E. Gardner and A. Zippelius, Europhysics Lett. **4**, 167 (1987).

[3] J.J. Arenzon and N. Lemke, J. Phys. A **27**, (1994) 5161

[4] M.A. Montemurro and F.A. Tamarit, Physica A **294**, 342 (2001).

[5] Watts, D. J., y S. H. Strogatz, Nature (London), **393**, 440 (1998).

[6] Barabási, A.-L., y R. Albert, Science, **286**, 509 (1999); Albert, R., H. Jeong, y A.-L. Barabási, Nature (London), **406**, 378 (2000).

[7] V. Eguiluz, D. Chialvo, G. Cecchi, M. Baliki and A.V. Apkarian, Phys. Rev. Lett. **92**, 018102 (2005).

[8] B.J. Kim, Phys. Rev. Lett. **93** (2004) 168701

[9] J. Altman and G.D. Gas, Nature **207**, 953 (1965); G. Cecchi, L. Petreanu, A. Alvarez-Buylla and M. Magnasco, J. Comp. Neurosci. **11**, 175 (2001). E. Gross, A.J. Reeves, M. Fallah, P. Tanapat, C.G. Gross and E. Fuchs, Proc. Natl. Acad. Sci. USA, **69**, 5263-7 (1999); S.A. Goldman, J. Neurobiol. **36**, 1444 (1990).

[10] D. Stauffer, A. Aharony, L. da Fontoura and J. Adler, Eur. Phys. J. B *32*, 395 (2003)

[11] L. da Fontoura and D. Stauffer, Physica A **330**, 37 (2003)

[12] G. Cziko, *Without Miracles. The universal selection theory and the second darwinian revolution*, a Bradford Book, MIT Press, Cambridge, USA (1995).