Nonequilibrium behavior in neural networks: criticality and optimal performance

J.J. Torres, S. Johnson, J.F. Mejias, S. de Franciscis, and J. Marro[1]

Institute Carlos I for Theoretical and Computational Physics, University of Granada,
E-18071, Granada Spain,
jtorres@onsager.ugr.es,
WWW home page: http://www.ugr.es/~jtorres

Abstract. We present a general theory which allows one to study the effects on emergent, cooperative behavior of a complex interplay between different dynamic processes that occur in actual systems at the neuron, synapse and network levels. We consider synaptic changes at different time scales from less than the millisecond to the scale of learning, and the possibility of finding a fraction of silent neurons. For some limits of interest, the fixed-point solutions or *memories* then loose stability and the system shows enhancement of its response to changing external stimuli for particular network topologies and *dynamical* memories. We observe at the edge of chaos that the network activity becomes critical in the sense that the relevant quantities show non–trivial, power–law distributions. We also describe the effect of activity–dependent synaptic processes on the network storage capacity.

1 Introduction

Recent research in neuroscience including both *in vivo* and *in vitro* experiments have demonstrated that synapses are more than simple communication lines among neurons, and that many different dynamic processes taking place in the synapses can influence and even determine different type of information processing in the brain [1]. Some of these mechanisms can occur on different time scales. For instance, on a time scale longer than the second (say days or years), synapses can be modified due to *learning.* This has been widely studied within a general theory of learning in attractor neural networks [2]. In addition to this, it has been described that fast synaptic fluctuations coupled with other mechanisms during the transmission of information seem to determine a large variety of computations in the brain [3, 4]. These fluctuations occur on very short (less than the millisecond) temporal scales, and they seem to have different causes. For instance, the stochasticity of the opening and closing of the neurotransmitter vesicles, variation in the postsynaptic response along the dendritic tree, which in turn has several sources (e.g., variations of the glutamate concentration in the synaptic cleft) and differences in the power released from different locations on the active zone of the synapses [5]. Together with these fast synaptic changes, it has been reported that the postsynaptic response is also affected by short-time

activity-dependent mechanisms which can decrease or increase the amount of available neurotransmitter and, consequently, the postsynaptic response is either depressed or facilitated [1, 6]. This type of synaptic plasticity is believed to be fundamental for the development and adaptation of the nervous system, and to be at the base of higher brain functions such as learning and memory.

In this paper we present an attempt towards a theoretical framework to study systematically the influence of synaptic changes on the collective properties of a neural network, where the network topology itself is also a variable to be considered. In particular, of special interest is to understand how these synaptic mechanisms for different network topologies affect the fixed points of the neural activity and their stability, which concerns memory, recall processes and sensibility to external stimuli.

2   Model and results

Let us consider N neurons – for simplicity assumed binary so that configurations are $\mathbf{S} \equiv \{s_i = \pm 1; i = 1, \dots, N\}$ – connected by synapses of intensity $w_{ij} = \overline{w}_{ij} z_j \;\; \forall i, j$. Here, $\overline{w}_{ij} \equiv 1/N \sum_{\mu=1}^{M} \Xi_i^\mu \Xi_j^\mu$ are fixed and determined in a previous *learning* process in which the M patterns of neural activity $\Xi^\mu \equiv \{\Xi_i^\mu = \pm 1; i = 1, \dots, N\}$, $\mu = 1 \dots M$, are *stored*. $\overline{w}_{ij}$ represents the maximal averaged synaptic conductance between the presynaptic neuron j and the postsynaptic neuron i, while, $z_j \in \mathbb{R}$ is a stochastic variable that influences these conductances and accounts for other synaptic dynamics than those associated to learning. For fixed $\mathbf{W} \equiv \{\overline{w}_{ij}\}$, the network state $\mathbf{A} = (\mathbf{S}, \mathbf{Z} \equiv \{z_i\})$ follows the probabilistic dynamics

$$\frac{\partial P_t(\mathbf{A})}{\partial t} = \sum_{\mathbf{A}'} \left[P_t(\mathbf{A}')c(\mathbf{A}' \to \mathbf{A}) - P_t(\mathbf{A})c(\mathbf{A} \to \mathbf{A}') \right], \tag{1}$$

where $c(\mathbf{A} \to \mathbf{A}') = p\, c^{\mathbf{Z}}(\mathbf{S} \to \mathbf{S}')\, \delta_{\mathbf{Z}, \mathbf{Z}'} + (1 - p)\, c^{\mathbf{S}}(\mathbf{Z} \to \mathbf{Z}')\, \delta_{\mathbf{S}, \mathbf{S}'}$ [7]. This amounts to assume that neurons ($\mathbf{S}$) change stochastically in time competing with a noisy dynamics of synapses ($\mathbf{Z}$), the latter with an *a priory* relative weight of $(1 - p)/p$ [8].

For $p = 1$, the model reduces to the Hopfield case, in which synapses are *quenched*, i.e., z_i is constant and independent of i, e.g., $z = 1$. This limit has been widely studied in the last decades [2]. More interesting is the case of $p \to 0$, which describes fast synaptic fluctuations. In this limit, one can uncouple the stochastic dynamics for neurons ($\mathbf{S}$) and the synaptic noise ($\mathbf{Z}$) using standard techniques [8]. It follows that neurons evolve as in the presence of a steady distribution for the noise $\mathbf{Z}$: If we write $P_t(\mathbf{A}) = P_t(\mathbf{Z}|\mathbf{S})P_t(\mathbf{S})$, where $P_t(\mathbf{Z}|\mathbf{S})$ stands for the conditional probability of $\mathbf{Z}$ given $\mathbf{S}$, one obtains from (1), after rescaling time $tp \to t$ and summing over $\mathbf{Z}$, that

$$\frac{\partial P_t(\mathbf{S})}{\partial t} = \sum_{\mathbf{S}'} \left\{ P_t(\mathbf{S}')\bar{c}[\mathbf{S}' \to \mathbf{S}] - P_t(\mathbf{S})\bar{c}[\mathbf{S} \to \mathbf{S}'] \right\}. \tag{2}$$

Here, $\bar{c}[\mathbf{S} \to \mathbf{S}'] \equiv \sum_{\mathbf{Z}} P^{\mathrm{st}}(\mathbf{Z}|\mathbf{S}) \, c^{\mathbf{Z}}[\mathbf{S} \to \mathbf{S}']$, and the stationary distribution for the noise is

$$P^{\mathrm{st}}(\mathbf{Z}|\mathbf{S}) = \frac{\sum_{\mathbf{Z}} c^{\mathbf{S}}[\mathbf{Z}' \to \mathbf{Z}] \, P^{\mathrm{st}}(\mathbf{Z}'|\mathbf{S})}{\sum_{\mathbf{Z}} c^{\mathbf{S}}[\mathbf{Z} \to \mathbf{Z}']}. \tag{3}$$

This expression involves an assumption on how synaptic noise depends on the overall neural activity. An interesting particular situation is to assume activity-dependent synaptic *noise* consistent with short-term synaptic depression and/or facilitation [6, 9]. That is, let us assume that $P^{\mathrm{st}}(\mathbf{Z}|\mathbf{S}) = \prod_j P(z_j|\mathbf{S})$ with

$$P(z_j|\mathbf{S}) = \zeta(\mathbf{m}) \, \delta(z_j - \varPhi) + [1 - \zeta(\mathbf{m})] \, \delta(z_j - 1). \tag{4}$$

Here, $\mathbf{m} = \mathbf{m}(\mathbf{S}) \equiv \left(m^1(\mathbf{S}), \ldots, m^M(\mathbf{S})\right)$ is the M-dimensional overlap vector, $m^\mu = N^{-1} \sum_i \varXi_i s_i$, and $\zeta(\mathbf{m})$ stands for a function of $\mathbf{m}$ to be determined. With this choice, the average over the distribution (4) of the noise variable is $\overline{z_j} \equiv \int z_j P(z_j|\mathbf{S}) dz_j = 1 - (1-\varPhi)\zeta(\mathbf{m})$ and the variance is $\sigma_z^2 = (1-\varPhi)^2 \zeta(\mathbf{m})[1 - \zeta(\mathbf{m})]$. Note that these two quantities depend on time for $\varPhi \neq 1$ through the overlap vector $\mathbf{m}$, which is a measure of the activity of the network. Moreover, the depression/facilitation effect in (4), namely $z_j = \varPhi > 0$ ($\varPhi \neq 1$), depends through the probability $\zeta(\mathbf{m})$ on the overlap vector, which is related to the net current arriving to postsynaptic neurons. Consequently, the non–local choice (4) introduces non–trivial correlations between synaptic noise and neural activity. One has a depressing (facilitating) effect for $\varPhi < (>)1$, and the trivial case $\varPhi = 1$ corresponds to the Hopfield model with quenched synapses. It is remarkable that, although the fast noise dynamics occurs at a very small time scale, the depressing or facilitating mechanism occurs on the time scale of the neural activity –via the coupling with the overlap vector through the function $\zeta(\mathbf{m})$.

The general model described by Eqs. (1-4) can be easily generalized to other cases of interest such as the possibility of having silent nodes in the network (every time the activity is updated). For instance, one may assume that the transition probabilities have the form

$$c^{\mathbf{Z}}[\mathbf{S} \to \mathbf{S}'] = \sum_{\mathbf{x}} p_n(\mathbf{x}) \prod_{i|x_i=1} \tau_n(s_i \to s_i'; \mathbf{Z}) \prod_{i|x_i=0} \delta_{s_i, s_i'}. \tag{5}$$

Here, $\mathbf{x}$ is an operational set of binary indexes fixed to 1 at n sites chosen at each time according to distribution $p_n(\mathbf{x})$, and fixed to zero at the other $N-n$ sites. The choice (5) simply states that one (only) updates simultaneously the selected n nodes. We also assume that the elementary rate $\tau_n(s_i \to s_i'; \mathbf{Z})$ depends on the factor $\beta s_i h_i$ where $\beta = T^{-1}$ is the inverse of the temperature and $h_i(\mathbf{S}, \mathbf{Z}) \equiv \sum_j \epsilon_{ij} \overline{w}_{ij} z_j s_j$ is the local field or synaptic current a particular neuron is receiving from its neighbors. Here $\epsilon_{ij} = \{1, 0\}$ is the adjacency or connectivity matrix, which describes the existence or not of interaction between i and j. This allows for the consideration of different network topologies. For any non-trivial topology, is convenient to consider the *local* overlap $m_j^\mu = \langle k \rangle^{-1} \sum_i \varXi_i s_i$, where $\langle k \rangle$ is the mean neuron degree or number of neighbors, which equals the *global* one only when all neurons are connected to each other.
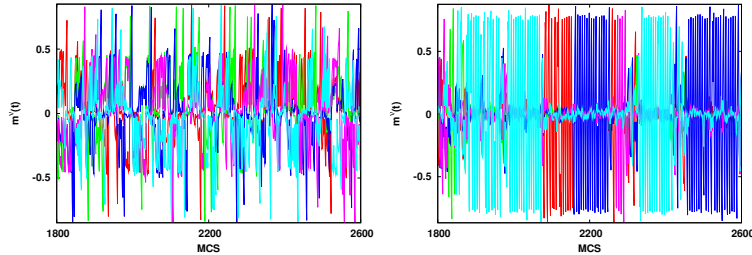
Fig. 1. Two of the different types of dynamic behaviour exhibited by the nonequilibrium neural network showing chaotic switching from one attractor to the other. Parameters are $N = 1600$, $M = 5$, $T = 0.01$, $\Phi = -0.8$ and $\rho = 0.45(0.5)$ for the left (right) simulation.

A main result is that the system shows quite qualitatively different behavior depending on the value of T and $\rho \equiv n/N$. In particular, one easily observes the familiar Hopfield phases of complete disorder and one in which the system may recover one of the stored patterns. More intriguing are cases such as those in Fig. 2, namely, dynamic phases in which the network activity chaotically switches among the stored patterns and antipatterns (i.e., negatives of the stored patterns).

Concerning topology, a given neuron i in biological networks is seldom or never connected to the rest, but rather only to a relatively small subset of neighbors – of cardinal $k_i = \sum_j \epsilon_{ij}$, which is i's *degree*. Even in the case of the worm *C. elegans*, with only about 300 neurons, it turns out that there is great disparity in the values of k, ranging from just one or two to a large portion of the network. Although little is yet known about the precise architecture of animal brains, a first approximation we can consider for neural systems – as is often done in other networks – is to assume that the synapses are placed randomly between the neurons, and focus only on the degree distribution, $p(k)$. In this approximation, known as the *configurations ensemble*, the expected value of ϵ_{ij}, given k_i and k_j, is $\overline{\epsilon_{ij}} = k_i k_j / (\langle k \rangle N)$. We will be particularly interested in *scale-free* distributions, $p(k) \sim k^{-\gamma}$, which appear to be quite ubiquitous in nature.

Standard mean-field analysis for $M = 1$ (a single stored pattern and its antipattern) [10] shows that there is a second order phase transition, from the memory phase to the chaotic one, with increasing temperature. The critical value for $\rho = 1$ is $T_c = \langle k^2 \rangle / (\langle k \rangle N)$, where the averages $\langle \cdot \rangle$ are over $p(k)$. Setting $T = 0$, we find there is also a transition, with decreasing Φ, from the memory phase in which either of the fixed points $m = \pm 1$ is stable, to one in which the system jumps chaotically between the two. The critical value of Φ at which this occurs is $\Phi_0 = 1 - \langle k \rangle^3 \langle k^3 \rangle^{-1}$. MC simulations show that, for non-zero T, chaotic windows open for $\Phi \lesssim \Phi_0$. They also tell us that these results are robust for larger values of M and qualitatively similar for $\rho < 1$.
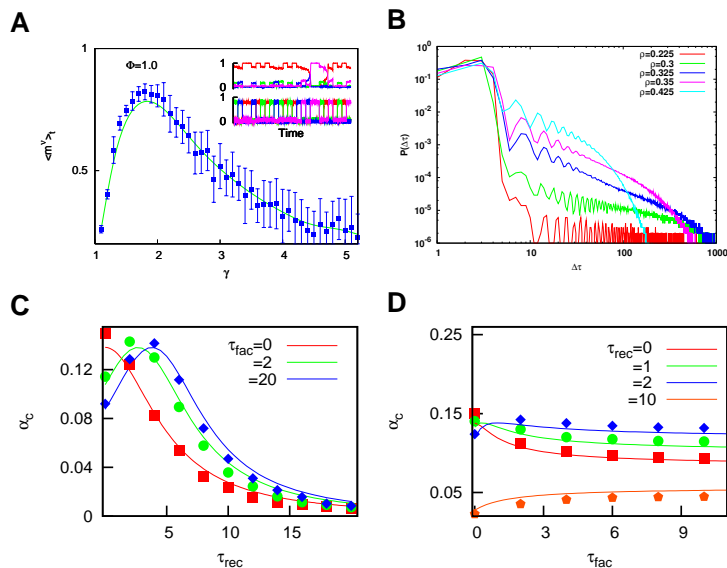
Fig. 2. (A) Network performance against γ for scale-free random topologies, with $\Phi = 1.0$ (static-synapses limit). Averages over 20 realisations, with stimulation every 50 MCS for 2000 MCS. Other parameters are $\delta = 5$, $M = 4$, $T = 2/N$, $\langle k \rangle = 20$, $N = 1600$, $\rho = 1$. Inset: sections of typical time series of m^ν, with $\nu = 1, ..., 4$ represented as different colours, for $\gamma = 4$ and $\Phi = 1$ (top) and $\Phi = 0.6$ (bottom). (B) Emergence of criticality in the model near the edge of chaos. Other parameter were $N = 800$, $M = 10$, $T = 0$, $\Phi = 0.75$ and $\Delta = 0.1$. (C) Critical storage capacity α_c as a function of τ_{rec}, for different values of τ_{fac}. (D) Critical storage capacity α_c as a function of τ_{fac}, for different values of τ_{rec}. In both panels, the results from numerical simulations (symbols) with ($N = 3000$) are supported by mean field predictions (lines).

This *edge of chaos* is particularly interesting, since it has been shown that this kind of transition can be optimal for certain magnitudes such as computational capacity [11] and dynamic range of sensitivity to stimuli [12]. To illustrate how this is also the case here, we store a set of M patterns and then *show* the system a randomly chosen one every certain number of time steps. This is done by changing the field at each node for just one MCS: $h_i \rightarrow h_i + \delta \xi_i^\nu$, as if a signal of intensity δ had been received. In general, when in the ordered phase, a high δ is required to destabilize the current pattern. However, close to the transition to chaos, the system becomes more sensitive and reacts appropriately. Within the chaotic regime, however, it does not then retain the pattern shown. We quantify performance as the temporal average $\langle m^\nu \rangle_{time}$, where ν is the pattern last shown. This is displayed in Fig. 2A. The figure shows that the performance exhibits a peak corresponding to the optimal scale-free topology, that for the case of low depression ($\Phi \rightarrow 1$) occurs around $p(k) \sim k^{-2}$. This is a consequence of the particular dependence of the critical parameters T_c and Φ_c with the moments of the degree distribution: the more heterogeneous the network, the more robust

the system will be to thermal fluctuations, and yet the less synaptic depression will be required to place the system at the edge of chaos. This optimization effect also is reflected by the emergence of power-law distributions of permanence times around the attractors, when the system in near the edge of chaos, as is depicted in Fig. 2B. Here, $P(\Delta\tau)$ is the probability that, throughout the time interval $(t', t' + \Delta\tau)$, a local field $h_i(t)$ is fluctuating above a baseline $h_0 = 0.1$ or below $h_0 = -0.1$, and averaged over all sites in the network. Optimal scale-free network topologies with exponent -2 could be related to the existence of a functional scale-free topology of the human brain during cognitive tasks with the same exponent [13]. It seems plausible, therefore, that the brain may acquire this kind of distribution – either structurally or functionally – in order to maximize its performance.

Finally we also report on the effect of synaptic processes coupled with network activity on maximum storage capacity of the network [14] via a phenomenological model of activity-dependent synapses (see [15] for details) which involves a competition between facilitating and depressing synaptic mechanisms. This model can be studied using our general theoretical framework assuming $P(z_j/\mathbf{S}) = \delta(z_j - \Phi_j(\mathbf{S}))$ with $\Phi_j(\mathbf{S})) = D_j(t, s_j)F_j(t, s_j)$ and $D_j(t, s_j)$, $F_j(t, s_j)$ representing dynamic variables for the depression and facilitation mechanisms varying in time-scales τ_{rec} and τ_{fac}, respectively. Maximum storage capacity is computed at $T = 0$ and in the memory phase. In practice, this implies considering only the steady state of D_j and F_j as a function of s_j and including them in the local field h_i. When the number of stored patterns is increased, the interference between patterns makes the attractors associated with these patterns loose their stability. Then, the maximum number of patterns, relative to the network size, namely $\alpha_c = M_c/N$, that the network can store and retrieve without interference is the maximum storage capacity. The behaviour of α_c as a function of the relevant synaptic parameters is depicted in Fig. 2C. In general, one has a non-monotonic dependence of storage capacity as a consequence of the competition of different synaptic processes with, *a priori*, opposite effects on the stability of attractors (e. g., depression tries to destabilize the memories whereas facilitation tries to keep the activity of the network in one particular attractor). This fact implies that activity-dependent synaptic changes are not only convenient for dynamic processing of information in actual neurons, as we have seen above, but an optimal balance between depression and facilitation effects is also necessary to have neural networks with good retrieval properties.

References

1. Abbott L.F. and Regehr W.G.: Synaptic computation. Nature, **431** (2004) 796–803.
2. Peretto, P. (1992). An Introduction to the modeling of neural networks. Cambridge University Press.
3. Allen C. and Stevens C.F.:An evaluation of causes for unreliability of synaptic transmission. Proc. Natl. Acad. Sci. USA **91** (1994) 10380–10383.

4. Zador A.: Impact of Synaptic Unreliability on the Information Transmitted by Spiking Neurons. J. Neurophysiol. **79** (1998) 1219–1229.
5. Franks K.M., Stevens C.F., and Sejnowski T.J.: Independent Sources of Quantal Variability at Single Glutamatergic Synapses. J. Neurosci. **23** (2003) 3186–3195.
6. Tsodyks M.V., Pawelzik K., Markram H.: Neural networks with dynamic synapses. Neural Comp. **10** (1998) 821–835.
7. Torres J.J., Garrido P.L., Marro J.: Neural networks with fast time-variation of synapses. J. Phys. A: Math. Gen. **30** (1997) 7801–7816.
8. Marro J., Dickman R., Nonequilibrium Phase Transitions in Lattice Models, Cambridge University Press, 1999.
9. Pantic L., Torres J.J., Kappen H.J., Gielen S.C.A.M.: Associative Memory with Dynamic Synapses. Neural Comp. **14** (2002) 2903–2923.
10. Johnson S., Marro J., Torres J.J.: Functional optimization in complex excitable networks. EPL **83** (2008) 46006.
11. Bertschinger N., Natschläger T.: 'Real-time computation at the edge of chaos in recurrent neural networks. Neural Comp. **16** (2004) 1413.
12. Assis V.R.V., Copelli M.: Dynamic range of hypercubic stochastic excitable media Phys. Rev. E**77** (2008) 011923.
13. Eguíluz V.M., Chialvo D.R., Cecchi G.A., Baliki M., Apkarian A.V., Scale-free brain functional networks, Phys. Rev. Lett. **95** (2005) 018102.
14. Torres J.J., Pantic L., Kappen H.J.: Storage capacity of attractor neural networks with depressing synapses. Phys. Rev. E. **66** (2002) 061910.
15. Mejias J.F., Torres J.J.: Maximum memory capacity on neural networks with short-term synaptic depression and facilitation. Neural Comput. **21** (2009) 851–871.