

# On the Role of Synaptic Depression in the Performance of Attractor Neural Networks

Joaquín J. Torres\*, Lovorka Pantic<sup>†</sup> and Hilbert J. Kappen<sup>†</sup>

\*Institute “Carlos I” for Theoretical and Computational Physics and Department of Electromagnetism and Material Physics, University of Granada, Fuentenueva s/n, E-18071 Granada, Spain.

<sup>†</sup>Department of Biophysics, University of Nijmegen, Geert Grooteplein, Nijmegen, The Netherlands

**Abstract.** Using a biologically motivated model of synaptic depression and within a mean-field approach, we examined the role of synaptic depression in the capacity of a binary neural network with  $N$  units to store and retrieve  $P$  patterns. In the limit of  $\alpha \equiv P/N \rightarrow 0$ , our results demonstrate the appearance of a novel phase characterized by quick transitions from one memory state to another. This phenomenon might reflect the flexibility of real neural systems to receive and respond to novel and changing external stimuli. In addition, we have computed the maximum storage capacity of such a network in the limit of  $\alpha \neq 0$  and  $T = 0$ . Supported by mean-field results and Monte Carlo simulations, we concluded that the critical storage capacity for effective retrieval of stable memory patterns decreases with the degree of the depression. Nevertheless, the storage of memories as *oscillatory* states will require a different definition of storage capacity. How such a new storage capacity depends on the synaptic depression is still an open question.

## INTRODUCTION

One of the interesting questions that arise in neural network modelling is how biological processes in real neurons at the cellular and subcellular level influence the network behavior and its ability to process information.

In most neural network studies so far, the synaptic connection between neurons has been modelled as a constant strength [1, 2], possibly subject to learning on a slow time scale. However, recently it has been reported that synaptic strength is a dynamic quantity that strongly depends on the presynaptic neural activity [3, 4]. Synaptic strength can decrease (depression) or increase (facilitation), depending on the type of synapse [3]. After learning, most synapses are depressing with a typical time constant of 10 – 20 *ms*. The synaptic strength recovers on the order of seconds. The molecular mechanism underlying this dynamics is the depletion of neurotransmitter vesicles due to presynaptic neural firing and their restoration on a larger time scale [4]. These synapses are found throughout the cortex as well as in the hippocampus [5].

There have been various studies of the effect of dynamic synapses on the information transfer in feed-forward neural network behavior, such as filtering of redundant noise [6], extraction of temporal patterns for speech recognition [7] and robust coincidence detection [8].

Only few studies, for instance [9, 10], have focused on the role of dynamical synapses

in recurrent neural networks, and in particular on associative memory [1]. In such networks, long-term storage of the memory patterns is produced by adjusting the strength of the synapses according to the Hebb rule. With such static synapses the network dynamics has fixed points (attractors) corresponding to the stored memory patterns [11]. The question one may ask is how the retrieval properties and the fixed points of the Hopfield network are affected by dynamical synapses.

In this paper, we show that the effect of depressing synapses is two-fold: 1) the memory capacity, in the sense of the stability of fixed points of the attractor dynamics, is strongly reduced and 2) the dynamics displays a new phase characterized by rapid switching among stored patterns. These results are observed not only in networks of binary neurons, but also of more biologically plausible integrate-and-fire neurons [10].

## THE MODEL

The model consists of a network of  $N$  binary neurons  $s_i = 1, 0$ . Each neuron follows a probabilistic dynamics which is given by:

$$\text{Prob}\{s_i(t+1) = 1\} = \frac{1}{2}[1 + \tanh\{2\beta h_i(t)\}]. \quad (1)$$

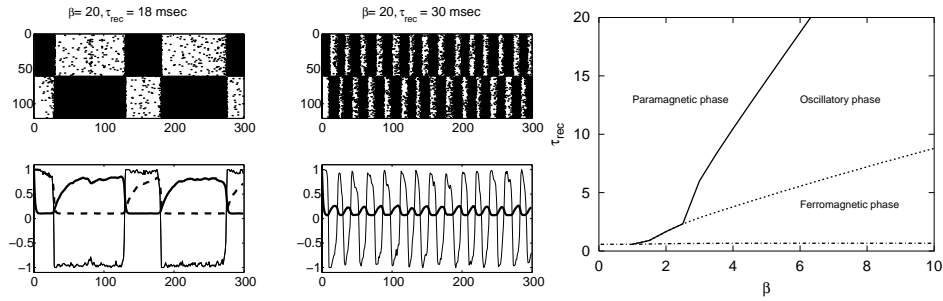
Here,  $h_i = \sum_j \omega_{ij} x_j s_j - \theta_i$  is the local field associated to neuron  $i$ ,  $\theta_i$  is the threshold for neuron  $i$  to fire and  $\beta = \frac{1}{T}$  represent the level of noise due to the stochastic synaptic activity. The first term of  $h_i$  is the total synaptic current arriving at neuron  $i$  with  $\omega_{ij}$  representing the static synaptic connection strength or *weights* between neurons  $j$  and  $i$  and  $x_j$  the depression variable associated with the pre-synaptic neuron [10]. The static weights are defined according to the standard covariance rule, that is:  $\omega_{ij} = \frac{1}{Nf(1-f)} \sum_{v=1}^P (\xi_i^v - f)(\xi_j^v - f)$ , where  $\xi_i^v = 0, 1$  are independent random bits with mean activity  $\langle \xi_i^v \rangle \equiv f = 1/2$ . Relatively good storage capacity is obtained when  $\theta_i = 0$ , when using  $s_i = \pm 1$  coding [11]. In  $\{1, 0\}$  code this corresponds to  $\theta_i = \frac{1}{2} \sum_j \omega_{ij}$ .

The depression variable  $x_j$  models the dynamic properties of the synapse [10] and follows the dynamics

$$x_i(t+1) = x_i(t) + \frac{1 - x_i(t)}{\tau_{rec}} - U x_i(t) s_i(t) \quad \forall i = 1, \dots, N. \quad (2)$$

where  $\tau_{rec}$  is the recovery time constant and  $U$  is the amount of resources in the active state released in each presynaptic event. For  $\tau_{rec} = 0$  we recover the standard static synapse ( $x_i = 1$ ). The synaptic model (2) is a simplified version of the phenomenological model by [4] (in the limit of  $\tau_{in} = 0$ ) in which each synapse is represented by a three-variable state whose dynamics is given by:

$$\begin{aligned} dx_i/dt &= z_i/\tau_{rec} - U x_i \delta(t - t_{sp}), \\ dy_i/dt &= -y_i/\tau_{in} + U x_i \delta(t - t_{sp}), \\ dz_i/dt &= y_i/\tau_{in} - z_i/\tau_{rec} \end{aligned} \quad (3)$$



**FIGURE 1.** Left: Periodic regimes appearing in a neural network of  $N = 120$  units with dynamic synapses, one stored pattern  $\mathbf{v}$  ( $\alpha \approx 0$ ) and for two different values of  $\tau_{rec}$  (Top panels). Evolution of  $m^{\mathbf{v}}$  and  $x_+^{\mathbf{v}}$  during the periodic regimes showed in the top panels (Bottom panels). Right: Phase plot ( $\beta, \tau_{rec}$ ) of attractor neural network with depressing synapses (for  $\alpha \approx 0$ ) displaying the retrieval (ferromagnetic), no retrieval (paramagnetic) and oscillatory phases.

with  $x_i + y_i + z_i = 1$ . Here,  $\delta(x)$  is the delta function and  $t_{sp}$  is the time at which a presynaptic spike occurs. The variables  $x_i$ ,  $y_i$  and  $z_i$  represent the fraction of synaptic resources in the recovered, active and inactive state, respectively. Then, the total synaptic current arriving to neuron  $i$  is proportional to the fraction of active resources, that is:  $I_i(t) = A \sum_j y_j$ .

## RESULTS

### *Switching among stored patterns*

Computer simulation of the model described by Eqs. (1-2) revealed that its behavior strongly depends on the parameters, such as the level of noise  $\beta$ , the number of patterns  $P$ , the amount of depression  $U$  and the recovery time constant  $\tau_{rec}$ . In the static case,  $\gamma \equiv \tau_{rec}U = 0$ , memories are attractors of the dynamics and are successfully retrieved when the noise is not too high ( $\beta$  sufficiently large) and the fraction of patterns  $\alpha = P/N$  is not too big (retrieval phase). Memories become unstable due to noise (paramagnetic phase) or due destructive interference between patterns, when  $\alpha$  becomes too large (spin-glass phase).

In the dynamic case, there exists an additional phase where patterns are stable for a relatively long times, but then rapidly switch to another pattern. This behavior is illustrated in Fig. 1.

Within a mean-field approach and for each stored pattern  $\mathbf{v}$ , we define the average of  $x_j$  for the active and inactive units in the pattern, that is  $x_+^{\mathbf{v}} \equiv \frac{2}{N} \sum_{j \in Act(\mathbf{v})} x_j$ , and  $x_-^{\mathbf{v}} \equiv \frac{2}{N} \sum_{j \notin Act(\mathbf{v})} x_j$ , respectively. Similarly, we define the mean activity for active and inactive units in pattern  $\mathbf{v}$ ,  $m_+^{\mathbf{v}} \equiv \frac{2}{N} \sum_{j \in Act(\mathbf{v})} s_j$  and  $m_-^{\mathbf{v}} \equiv \frac{2}{N} \sum_{j \notin Act(\mathbf{v})} s_j$ , respectively.

Then, in the thermodynamic limit ( $N \rightarrow \infty$ ), the corresponding local fields are given by:

$$h_i = \frac{1}{2} \sum_{v=1}^P \varepsilon_i^v [x_+^v m_+^v - x_-^v m_-^v], \quad (4)$$

where  $\varepsilon_i^v \equiv 2\xi_i^v - 1$ , and  $m^v \equiv \frac{2}{N} \sum_i \varepsilon_i^v s_i = m_+^v - m_-^v$  is the overlap of the network activity with a stored pattern  $v$ .

When the number of patterns  $P$  is small and the network is large ( $\alpha \approx 0$ ), we can assume that the network has a large overlap with pattern  $v = 1$  and neglect the sum over  $P > 1$  in (4). Equations (1) and (2) then reduce to the map:

$$m_{\pm}^1(t+1) = \frac{1}{2} \{1 \pm \tanh[\beta(x_{\pm}^1(t)m_{\pm}^1(t) - x_{\mp}^1(t) - m_{\pm}^1(t))]\}, \quad (5)$$

$$x_{\pm}^1(t+1) = x_{\pm}^1(t) + \left[ \frac{1 - x_{\pm}^1(t)}{\tau_{rec}} - U x_{\pm}^1 m_{\pm}^1 \right]. \quad (6)$$

Linear stability analysis of (5-6) reveals that there exists two stable fixed-point regimes corresponding to *ferromagnetic* and *paramagnetic* type of behavior. For small  $\beta$  there exists one stable paramagnetic fixed point:  $m_{\pm}^1 = \frac{1}{2}$ . For large  $\beta$  there exists two stable ferromagnetic fixed point solutions  $m_{\pm}^1 \approx 1, m_{\mp}^1 \approx 0$ . In both cases,  $x_{\pm}^1 = \frac{1}{1 + \gamma m_{\pm}^1}$ . For intermediate values of  $\beta$  and for large enough  $\tau_{rec}$ , neither of these solutions is stable and the network displays oscillatory behavior (see Fig. 1 for a typical behavior and the phase portrait).

Preliminary simulations confirm that the oscillatory phase remains for  $\alpha \neq 0$  (see Fig. 2 left). In this case, the oscillatory behavior is more complex due to the interference between several patterns and the presence of spin-glass states.

### Storage capacity

We analyzed the storage capacity of the network when we store an infinity number of patterns ( $\alpha \neq 0$  and finite) and for all possible depression strengths and recovery time constants. Following the standard approach of [2] (AGS) and assuming the network has a macroscopic overlap with one pattern (*condensed pattern*,  $v = 1$ ) in the steady state, and the remaining  $P - 1$  overlaps are  $O(\frac{1}{\sqrt{N}})$ , the generalization of (5-6) for  $\alpha$  finite is:

$$\begin{aligned} u &= \int \frac{dz}{\sqrt{2\pi}} e^{-z^2/2} \tanh \left[ \beta \left( \frac{4u}{\gamma^2(1-u^2) + 4\gamma + 4} + \sqrt{\alpha} r z \right) \right], \\ q &= \int \frac{dz}{\sqrt{2\pi}} e^{-z^2/2} \tanh^2 \left[ \beta \left( \frac{4u}{\gamma^2(1-u^2) + 4\gamma + 4} + \sqrt{\alpha} r z \right) \right], \\ r &= \frac{q}{[1 - \beta(1-q)]^2}, \end{aligned} \quad (7)$$

where  $u \equiv 2m_+^1 - 1$ ,  $q$  is the *spin-glass* order parameter [2] and  $r \equiv \frac{1}{\alpha} \sum_{\mu \neq 1} (x_+^\mu m_+^\mu - x_-^\mu m_-^\mu)^2$ .<sup>1</sup> The system (7) reduces to the AGS equations in the limit of  $\gamma \rightarrow 0$ , that is, in the case of non-depressed or *static* synapses.

In general, the system (7) must be solved numerically. In the limit  $T \rightarrow 0$  it is easy to compute the critical value of  $P/N$ , namely  $\alpha_c$ , at which nontrivial ferromagnetic solutions (*Mattis states*) appear.  $\alpha_c$  gives the maximum number of patterns that the system is able to retrieve. For  $\beta \rightarrow \infty$ , we use the standard approximation  $\int \frac{dz}{\sqrt{2\pi}} e^{-z^2/2} (1 - \tanh^2 \beta[az + b]) \approx \sqrt{\frac{2}{\pi}} \frac{1}{a\beta} e^{-b^2/2a^2}$  and  $\int \frac{dz}{\sqrt{2\pi}} e^{-z^2/2} \tanh \beta[az + b] \approx \operatorname{erf}\left(\frac{b}{\sqrt{2a}}\right)$ . Then, introducing a new variable  $y = f(u, \gamma)/\sqrt{2\alpha r}$  with  $f(u, \gamma) \equiv \frac{4u}{\gamma^2(1-u^2)+4\gamma+4}$ , the system (7) reduces to

$$y \left( \sqrt{2\alpha} + \frac{2}{\sqrt{\pi}} e^{-y^2} \right) = f[\operatorname{erf}(y), \gamma]. \quad (8)$$

One can check that  $u = 0$  ( $y = 0$ ) is a solution of (8). The largest  $\alpha$  at which a non-zero solution (stable ferromagnetic state) appears defines the critical storage capacity  $\alpha_c$  at  $T = 0$  as a function of the degree of synaptic depression  $\gamma$ . In Fig. 2 right we plot the line  $\alpha_c(\gamma)$  (solid line) which predicts a decrease of  $\alpha_c$  as the degree of depression is increased. For  $\gamma \rightarrow 0$  (non-depressed synapses), we recover the well known critical storage capacity for static synapses, that is  $\alpha_c \approx 0.138$ . For  $\gamma \rightarrow \infty$  our mean-field theory predicts  $\alpha_c \rightarrow 0$ .

Numerical simulation of the network confirms our main mean-field results [12]. In general, the value of  $\alpha_c(\gamma)$  obtained in numerical simulations is slightly higher than the mean-field prediction (see data points in Fig. 2 right). This discrepancy must be attributed to the relatively small network sizes used in our numerical studies, possible non-linear finite size scaling effects at larger  $N$ , and the criterion we use to define  $\alpha_c(\gamma)$  for the destabilization of memory pattern. In addition, our mean field approach uses the replica symmetric ansatz, which is well-known to underestimate the exact value of  $\alpha_c$  for static synapses [2].

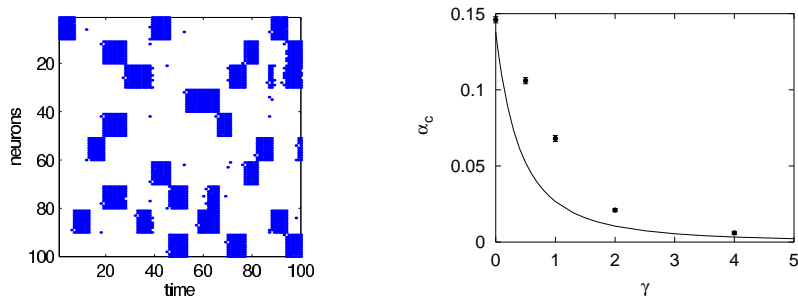
## DISCUSSION

We conclude, that attractor neural networks with dynamical (depressing) synapses display in addition to a ferromagnetic phase an oscillatory phase where the network activity switches between stored memories. None of these memories are fixed points, but rather meta-stable states. This behavior is also observed at high memory loading ( $\alpha \neq 0$ ). Therefore, the disappearance of memories as fixed points does not imply the disappearance of memory from these networks.

In addition, we shown that the traditional definition of storage capacity of stable memories (ferromagnetic states) decreases with the degree of depression. However, the storage of memories as meta-stable states will require a different definition of storage

---

<sup>1</sup> For a detailed derivation of (7) see [12].



**FIGURE 2.** Left: Raster plot of a network with 100 neurons and  $P = 10$  and non-orthogonal patterns ( $f = 0.1$ ). In the  $i$ th pattern neuron  $10(i - 1) + 1, \dots, 10i$  are active. Right: Critical storage capacity  $\alpha_c$  vs  $\gamma = \tau_{rect}$  in a neural network with depressing synapses at  $T = 0$ . Solid line is the mean-field result from Eqs. 7. Numerical data points correspond to the finite-size scaling analysis ( $N = 400, 800, 1600$ ) of the Monte Carlo simulations [12].

capacity. How such a new storage capacity depends on the dynamics of the synapses is an open question.

Simulations of a network consisting of integrate-and-fire neurons with dynamic synapses show the same switching capability among the memory states [10]. In order to determine the relevance of this behavior for neurobiology, further analyzes of the switching phenomenon will require more realistic network models as well as more realistic learning rules.

## ACKNOWLEDGMENTS

J.J.T. acknowledges support from “Ministerio de Ciencia y Tecnología” and FEDER (“Ramón y Cajal” contract and project no BFM2001-2841) and the Dutch Foundation for Neural Networks (SNN).

## REFERENCES

1. Hopfield, J., *Proc. Nat. Acad. Sci. USA*, **79**, 2554–2558 (1982).
2. Amit, D., Gutfreund, H., and Sompolinsky, H., *Ann. of Phys.*, **173**, 30–67 (1987).
3. Abbott, L., Varela, J., Sen, K., and Nelson, S., *Science*, **275**, 220–224 (1997).
4. Tsodyks, M., and Markram, H., *Proc. Natl. Acad. Sci. USA*, **94**, 719–723 (1997).
5. Dobrunz, L., and Stevens, C., *Neuron*, **18**, 995–1008 (1997).
6. Natschläger, T., Maass, W., and Zador, A., *Network: Comput. Neural Syst.*, **12**, 75–87 (2001).
7. Liaw, J. S., and Berger, T. W., *Hippocampus*, **6**, 591–600 (1996).
8. Pantic, L., Torres, J., and Kappen, H., *Neurocomputing*, **38-40**, 285–291 (2001).
9. Bibitchkov, D., Herrmann, J. M., and Geisel, T., *Network: Comput. Neural Syst.*, **13**, 115–129 (2002).
10. Pantic, L., Torres, J., Kappen, H., and Gielen, S., *Neural Computation*, *in press* (2002).
11. Amit, D., *Modeling brain function*, Cambridge University Press, Cambridge, 1989.
12. Torres, J., Pantic, L., and Kappen, H., *Phys. Rev. E*, *in press* (2002).