

To appear in Neural Computation 2002

Associative memory with dynamic synapses

Lovorka Pantic, Joaquín J. Torres*, Hilbert J. Kappen,
and Stan C.A.M. Gielen

Department of Biophysics, University of Nijmegen,
Geert Grooteplein Noord 21, 6525 EZ Nijmegen,

The Netherlands

April 18, 2002

Abstract

We have examined a role of dynamic synapses in the stochastic Hopfield-like network behaviour. Our results demonstrate an

*Present address: Instituto Carlos I de Física Teórica y Computacional and Departamento de Electromagnetismo y Física de la Materia, Facultad de Ciencias, Universidad de Granada, E-18071 Granada, Spain.

appearance of a novel phase characterized by quick transitions from one memory state to another. The network is able to retrieve memorized patterns corresponding to classical ferromagnetic states but switches between memorized patterns with an intermittent type of behaviour. This phenomenon might reflect the flexibility of real neural systems and their readiness to receive and respond to novel and changing external stimuli.

1 Introduction

Recent experimental studies report that synaptic transmission properties of cortical cells are strongly influenced by recent presynaptic activity (Abbott et al., 1997; Tsodyks and Markram, 1997). The postsynaptic response is activity-dependent and can decrease (in case of synaptic depression) or increase (in case of synaptic facilitation) under repeated stimulation (Tsodyks and Markram, 1997). For some cortical synapses it was found (Markram and Tsodyks, 1996) that after induction of long-term potentiation (LTP), the temporal synaptic response was not uniformly increased. Instead, the amplitude of the initial postsynaptic potential was potentiated whereas the

steady-state synaptic response was unaffected by LTP (synaptic redistribution).

These findings affect the transmission properties of single neurons as well as the network functioning and behavior. Previous studies on the effects of synaptic depression on network behavior have focused on both feed-forward and recurrent networks. For feed-forward networks, some authors have studied a role of activity-dependent synapses in a supervised learning paradigm (Natschlger et al., 2001). They showed that even with a single hidden layer of binary neurons such networks can approximate a large class of nonlinear filters. Another study (Liaw and Berger, 1996) demonstrated that dynamically changing synaptic strength allows the extraction of statistically significant features from noisy and variable temporal patterns. An application was presented where a simple neural network with dynamic synapses was able to perform speech recognition from unprocessed noisy waveforms of words spoken by multiple speakers.

For recurrent networks, a number of theoretical and numerical studies revealed that a large population of excitatory neurons with depressing synapses can exhibit complex regimes of activity (Bressloff, 1999; Kistler and Hemmen, 1999; Senn et al., 1996; Tsodyks et al., 1998; Tsodyks et al., 2000).

In (Tsodyks et al., 2000) it was demonstrated that such networks exhibit short intervals of highly synchronous activity (population bursts) intermittent with long periods of asynchronous activity resembling the behavior observed in neurons throughout the cortex. It was proposed (Senn et al., 1996; Senn et al., 1998) that synaptic depression may also serve as a mechanism which can provoke rhythmic activity and, therefore, can have an important role in central pattern generation.

The effects of synaptic dynamics on some complex neural functions such as associative memory, for instance, are not yet fully established and only few studies (Bibitchkov et al., 2002) have focused on this important issue. The conventional way for model studies of associative memory functions is based on the view that recurrent (attractor) neural networks may capture some basic characteristics of associative memory (Amit, 1989; Miyashita and Chang, 1988). In such networks, long-term storage of the memory patterns is enabled by synaptic connections which are adjusted (according to the Hebb rule) such that the attractors (fixed points) of the network dynamics represent the stored memories (Hopfield, 1982). The question one may ask is how the retrieval properties and the fixed points of the Hopfield network are affected by the experimental findings showing use-dependent weakening of the

synaptic connections. In (Bibitchkov et al., 2002), it was reported that short-term synaptic depression does not affect the fixed points of the network but rather reduces the stability of the patterns and the storage capacity. These authors also showed that during external stimulation, depression mechanism enables an easier switching between patterns and a better storage of temporal sequences.

In this study we aim to extend the previous work and to explore the role of noise as well as the effect of the synaptic recovery process in pattern retrieval. We found that a network with depressing synapses is sensitive to noise and displays rapid switching among stored memories. The switching behaviour is observed for overlapping as well as for non-overlapping patterns. It exists for the network consisting of simple binary units and also for the network with the integrate-and-fire neurons. The transitions among memories start to occur when the network dynamics, described by an iterative (mean-field) map, undergoes an intermittent behavior. The memory states lose absolute stability and the system, instead of residing at one of the fixed-point attractors, reveals another type of dynamics. Linear stability analysis as well as numerical simulations reveal that in addition to traditional ferromagnetic (memory) and paramagnetic (noisy) fixed-point regimes there exists a limit-cycle phase

which enables transitions among the stored memories.

2 Models

The classical Hopfield-like neural network consists of N fully connected binary neurons. The neurons are regarded as stochastic two-state devices with the state-variable s taking values 1 (firing) or 0 (no firing) according to the probabilistic rule given by:

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

Parameter $\beta = \frac{1}{T}$ represents the level of noise caused by synaptic stochasticity or other fluctuations in neuron functioning. The local field h_i is the net input to neuron i given by $h_i = \sum_j \omega_{ij}^0 s_j - \theta_i$, where ω_{ij}^0 represents the synaptic connection strength (weights) from neuron j to neuron i . The updating of all neurons can be carried out synchronously (Little dynamics) at each time step or asynchronously (Glauber dynamics), updating them one at a time. The connectivity matrix ω_{ij}^0 is defined according to the standard covariance rule

$$w_{ij}^0 = \frac{1}{Nf(1-f)} \sum_{\mu=1}^P (\xi_i^\mu - f)(\xi_j^\mu - f) \quad (2)$$

which represents an optimal learning rule for associative memory networks (Dayan and Willshaw, 1991; Palm and Sommer, 1996). The ξ^μ 's ($\mu = 1, \dots, P$) denote P patterns to be memorized (stored) and later, to be retrieved by the network. All thresholds θ_i are set to zero (Peretto, 1992). The variable f represents the mean level of activity of the network for the P patterns. It has been shown (Amit, 1989) that the memorized patterns are stable fixed points of the network dynamics as long as $\alpha \equiv \frac{P}{N}$ is small and as long as the noise level is not too high (Amit, 1989; Peretto, 1992) (ferromagnetic phase). When $\beta < 1$, the noise is stronger than the connectivity and memory is lost (paramagnetic phase). When α is too large, the interference between patterns destroys pattern stability and causes the spin glass phase to dominate (Amit, 1989).

As a model of the synaptic depression we use the phenomenological model introduced in (Tsodyks and Markram, 1997). This model assumes that synaptic resources are available in a fraction of recovered (x), active (y) and inactive (z) states and that the total amount of resources (neurotransmitter) in the recycling process is constant, which gives $x + y + z = 1$. Each action potential (AP) activates a fraction (U) of the resources in the recovered state which (instantaneously) becomes active (during transmission),

then becomes inactivated with a time constant (τ_{in}) of a few milliseconds and finally recovers with a time constant (τ_{rec}) that can vary from tens of milliseconds to seconds, depending on the type of the neuron and the type of synapse. The synaptic current is proportional to the fraction of active resources $I(t) = Ay(t)$ where A determines the maximal response of the synapses. τ_{in} is used to model the α -function like response of the synapse. The model can be further simplified since the inactivation time constant τ_{in} is much shorter than the recovery constant τ_{rec} (Tsodyks et al., 1998). The detailed synaptic response can be ignored and we can eliminate the variable y from the synaptic dynamics. Thus, we assume that the depressing synapse can be described by the dynamics of the recovered resources only. Since the states of the binary neurons are updated at discrete time steps we use a discretized equation

$$x(t + \delta t) = x(t) + \delta t \left(\frac{1 - x(t)}{\tau_{rec}} - Ux(t)s(t) \right) \quad (3)$$

instead of the original differential equation (Tsodyks and Markram, 1997; Tsodyks et al., 1998), with δt denoting the time step of discretization (updating).

To incorporate the synaptic dynamics into the network, we consider that during memory retrieval, the strength of synaptic connections (weights)

changes in time according to the depressing mechanism. The dynamic synaptic strength $\omega_{ij}(t)$ is assumed to be proportional to the fraction of recovered resources $x_j(t)$ and the static Hebbian term w_{ij}^0 , that is, $w_{ij}(t) = w_{ij}^0 x_j(t)$ ¹. The local fields are then given by $h_i = \sum_j w_{ij}^0 x_j s_j$. The states of all the neurons are updated synchronously at time steps of $\delta t = 1$ msec and together define the consecutive network states. The overlap of the instantaneous network state with the memorized pattern μ , i.e. $m^\mu \equiv \frac{1}{Nf(1-f)} \sum_{i=1}^N (\xi_i^\mu - f) s_i$, provides an insight into the network dynamics and measures similarity of the actual state and the stored patterns.

3 Simulation results

3.1 Network with binary units

We analyze the behaviour of a network consisting of N fully connected binary neurons for a small number of patterns. We first consider the case of a single pattern with $f = \frac{1}{2}$. The pattern is stored according to the rule (2) and the states of the neurons are governed by the equations (1) and (3).

¹It is worth mentioning that under this assumption, interactions between the neurons are no longer symmetric.

For static synapses it is well known that the network exhibits two types of behaviour. The pattern (anti-pattern) is either permanently retrieved (ferromagnetic phase) or the retrieval is not possible (paramagnetic phase). For depressing synapses, the pattern is only stable for a certain period of time after which it disappears (cf. figure 1 (top)). All efferent (outgoing) synapses from active units are permanently depressed during this period, i.e. their average level of recovery $x_+ \equiv \frac{1}{Nf} \sum_{j \in Act} x_j$ decreases to a low level (cf. figure 1 (bottom)). At the same time, all efferent synapses from inactive units recover, i.e. $x_- \equiv \frac{1}{N(1-f)} \sum_{j \notin Act} x_j$ increases. Due to noise, at some moment of time the network switches to the attractor formed by the recovered synapses (anti-pattern), see figure 1 (top). The anti-pattern is now temporarily active, then it disappears while the original pattern being instantaneously reactivated and so forth. This process repeats itself and figure 1 (bottom) shows the repetitive behavior of the x_+ , x_- and overlap m^1 for two values of τ_{rec} . For smaller values of τ_{rec} , the duration of the pattern (anti-pattern) activity is longer than for larger values of τ_{rec} for which the switching between attractor states occurs more frequently.

In case of several orthogonal patterns ² whose active units overlap (fig. 2

² $\frac{1}{N} \sum_{i=1}^N (\xi_i^\mu - f)(\xi_i^\nu - f) = 0$ when $\mu \neq \nu$

(top)) the network also displays switching behaviour. Panels showing overlaps with $P = 6$ patterns indicate that the network first recalls pattern #6 and its anti-pattern, alternately, (as shown in the sixth panel of the figure), then switches to pattern #3 and its anti-pattern (third panel) and then retrieves pattern #5 and its corresponding anti-pattern (fifth panel). Switching that occurs on the fast time scale between the pattern and its anti-pattern is similar to the behavior shown in figure 1 (limit-cycle). In addition there exists a switching on the slower time scale among the different patterns. The switching shown in fig. 2 is more complex than the switching previously considered (fig. 1) since one limit-cycle switches to another pattern or to another limit-cycle. This motion might be related to the behaviour that can occur for several co-existing attractive states.

A similar behaviour is also present in the case when several patterns with non-overlapping active units are stored into the network. The upper panel of fig. 3 shows the activity of the network as a function of time and the middle panel shows the level of recovery x_j for neuron j ($j = 1, \dots, N$). All synapses from temporarily active units become depressed while synapses from inactive units are either recovered or are recovering (cf. figure 3 (middle)). Random synchronous activity of some units from an inactive group (which occurs due

to noise) will cause an increase of their local fields and further excitation of the other units in the same group (associative mechanisms). Meanwhile, the local fields of the neurons from the currently active group become suppressed and activity of that group is inhibited. As a result of this process, the network switches to other (mixture) states and so forth.

3.2 Network with integrate-and-fire units

We will explore now the switching behaviour for the more realistic network with the integrate-and-fire (IF) neurons. We consider $N = 200$ fully connected neurons whose states are modeled by the membrane potential $V_i(t)$ ($i = 1, \dots, N$) according to the dynamics:

$$\tau_m dV_i(t)/dt = -V_i(t) + RI_i^{syn}(t) + \xi(t). \quad (4)$$

The membrane time constant and the membrane resistance are fixed to $\tau_m = 20 \text{ msec}$ and $R = 100 \text{ M}\Omega$, respectively. $\xi(t)$ represents Gaussian white noise with autocorrelation $\langle \xi(t)\xi(t') \rangle = D\delta(t - t')$. The synaptic input to cell i from other neurons is given by $I_i^{syn} = \eta A \sum_{j=1}^N w_{ij}^0 y_j$, where Ay_j represents the synaptic current I_j of the neuron j (see Models), w_{ij}^0 denotes the connectivity matrix given by 2 and η is a scaling factor. Whenever the membrane potential is equal to the threshold value ($V_{th} = 10 \text{ mV}$)

an action potential is emitted and the potential is reset to zero with the refractory period of $\tau_{ref} = 1 \text{ msec}$. To enable easier observation we store $P = 5$ patterns whose active bits do not overlap. Figure 4 shows the network behaviour during the pattern retrieval. For some fixed noise level and depending on the value of the time constant for recovery τ_{rec} , the network either permanently recalls one of the stored patterns ($\tau_{rec} = 0 \text{ msec}$ - standard behaviour of the network with constant weights), or switches among patterns ($\tau_{rec} = 300 \text{ msec}$ and $\tau_{rec} = 800 \text{ msec}$), or exhibits a paramagnetic-type of the behaviour ($\tau_{rec} = 3000 \text{ msec}$) when retrieval of the patterns is not possible. We conclude that the effects of synaptic depression on the network level are qualitatively the same for the network with integrate-and-fire units as it is for the network with the binary neurons. The most interesting feature is that synaptic depression destabilizes the stored memories and enables noise-induced transitions from one pattern to another. It has been demonstrated (Mueller and Herz, 1999) that the associative properties of a network with the IF neurons or with binary neurons are similar. Our results extend this observation for network with depressing synapses.

4 Mean-field calculations

To further investigate the switching phenomenon we have analyzed the behavior of the network with the binary units within the mean-field framework. For simplicity, we consider the case $\alpha \rightarrow 0$. When the network is close to pattern $\mu = 1$ the local field is given by :

$$\begin{aligned}
 h_i &= \sum_j w_{ij}^0 x_j \langle s_j \rangle \\
 &= \frac{1}{Nf(1-f)} \sum_j \sum_{\mu} (\xi_i^{\mu} - f)(\xi_j^{\mu} - f) x_j \langle s_j \rangle \\
 &\approx \frac{1}{Nf(1-f)} \sum_j (\xi_i^1 - f)(\xi_j^1 - f) x_j \langle s_j \rangle
 \end{aligned} \tag{5}$$

where we have ignored contributions from the other patterns. This term scales as $\sqrt{\frac{P}{N}}$ for random patterns. The sum \sum_j in the last expression can be split in two terms:

$$h_i = \frac{1}{Nf(1-f)} (\xi_i^1 - f) \left[\sum_{j \in Act} (\xi_j^1 - f) x_j \langle s_j \rangle + \sum_{j \notin Act} (\xi_j^1 - f) x_j \langle s_j \rangle \right] \tag{6}$$

where $j \in Act$ ($j \notin Act$) denote the active (inactive) neurons in pattern 1. We denote by x_+ the mean level of recovery and by $m_+ \equiv \frac{1}{Nf} \sum_{j \in Act} \langle s_j \rangle$ the mean activity of the active units. For the inactive units we have x_- and $m_- \equiv \frac{1}{N(1-f)} \sum_{j \notin Act} \langle s_j \rangle$. With these definitions we have $m^1 = m_+ - m_-$.

The local field is now given by:

$$h_i = (\xi_i^1 - f)(x_+m_+ - x_-m_-) \quad (7)$$

Hence, for m_+ we have

$$m_+(t+1) = \frac{1}{2} \{1 + \tanh [2\beta(1-f)(x_+(t)m_+(t) - x_-(t)m_-(t))]\} \quad (8)$$

and for m_-

$$m_-(t+1) = \frac{1}{2} \{1 - \tanh [2\beta f(x_+(t)m_+(t) - x_-(t)m_-(t))]\} \quad (9)$$

The dynamics of variables x_+ and x_- is governed by the equations

$$x_+(t+1) = x_+(t) + \left[\frac{1 - x_+(t)}{\tau_{rec}} - Ux_+(t)m_+(t) \right] \quad (10)$$

$$x_-(t+1) = x_-(t) + \left[\frac{1 - x_-(t)}{\tau_{rec}} - Ux_-(t)m_-(t) \right] \quad (11)$$

which together with equations (8) and (9) represent the four dimensional iterative map that approximately describes the network dynamics.

5 Stability analysis

We analyze the stability of the four dimensional map for the case $f = \frac{1}{2}$. The fixed points are given by the equations:

$$m_+ = \frac{1}{2} \{1 + \tanh [\beta(x_+m_+ - x_-m_-)]\}$$

$$\begin{aligned}
m_- &= \frac{1}{2} \{1 - \tanh[\beta(x_+ m_+ - x_- m_-)]\} \\
x_+ &= \frac{1}{1 + \gamma m_+} \\
x_- &= \frac{1}{1 + \gamma m_-}
\end{aligned} \tag{12}$$

with $\gamma = U\tau_{rec}$. Let us denote by

$$\vec{x}' = \vec{g}(\vec{x})$$

the iterative map with $\vec{x} = (m_+, m_-, x_+, x_-)$ and by

$$\vec{x} = \vec{x}^*$$

the fixed-point solution. Linear stability analysis demands linearization of $\vec{g}(\vec{x})$ near \vec{x}^* :

$$x'_i = g_i(\vec{x}^*) + \sum_j \frac{\partial g_i(\vec{x}^*)}{\partial x_j} (x_j - x_j^*) + \dots \tag{13}$$

The eigenvalues λ_i of the matrix $D \equiv \frac{\partial g_i(\vec{x}^*)}{\partial x_j}$ indicate the stability of the fixed-point \vec{x}^* . If $|\lambda_i| < 1 \forall i$, the solution is stable. In figures 5 and 6 we illustrate the stability of the system (8-11) for two different parameter sets. The stability matrix D and the eigenvalues are given in Appendix. Figure 5 shows that there are two critical (bifurcation) values of τ_{rec} (τ_1 and τ_2) where $|\lambda|_{max}$ crosses the value 1³. For $\tau_{rec} < \tau_1$, the system has

³The crossing which occurs for $\tau_{rec} < 1$ is not relevant and represents an approximation artifact since our map is not valid in that range.

three fixed points where two of them are stable and correspond to memory (ferromagnetic) solutions. As τ_{rec} increases, the fixed points approach each other and coalesce near $\tau_{rec} = \tau_1$. The figure shows that near this point the system undergoes a qualitative change in behavior. For $\tau_1 < \tau_{rec} < \tau_2$ stable oscillations occur (limit cycle attractor). The amplitude of oscillations (black and white circles in the figure) gradually decreases and finally ceases at τ_2 . At this point, the oscillatory behavior is replaced by a stable ($|\lambda|_{max} < 1$) fixed point which represents the no-memory (paramagnetic) solution (Hopf bifurcation). In sum, we see that the network dynamics, in addition to the fixed point (ferromagnetic and paramagnetic) attractors, reveals the limit cycle attractor which enables a periodic kind of behavior. In figure 6 we present the stability of the map for a larger value of β . It can be noticed that for $\tau_{rec} < \tau_1$, there exist 5 fixed points instead of 3 fixed points. The additional fixed points are unstable (saddle) and near τ_1 they coalesce with the stable fixed points. Upon further increase of τ_{rec} , both the stable points and unstable (saddle) points disappear (a saddle-node type of bifurcation). The network dynamics is illustrated in figure 7 where the fixed points are given by the intersections of the diagonal and the map. The intersection of the map with the diagonal at the value $m_+ = 0.5$ represents an unstable fixed

point. The values $m_+ \in \{0, 1\}$ are stable fixed points in the ferromagnetic phase. However, in the oscillatory regime they are no longer fixed points. The consecutive states of the network at equal time intervals (black dots in the figure) show that for $\tau_{rec} \gg \tau_1$ (upper left panel) the system reveals a limit-cycle behavior. For smaller values of τ_{rec} the system spends more time near one of the ferromagnetic fixed points. When τ_{rec} is almost equal to τ_1 the system resides for a long time near one of the ferromagnetic fixed points. In this condition the lower panels in Fig. 7 show the existence of two narrow “corridors” between the diagonal and the map. If the state of the system is in the neighborhood of these corridors, the system resides there for a long time before it escapes and approaches another corridor. Such a non-trivial dynamics resembles intermittent behavior (Hao, 1989) and represents in fact the beginning of a periodic motion (cycle). Figure 8 shows that the period of oscillations for the critical τ_1 is very large. This is also confirmed in numerical simulations of the stochastic system (cf. fig.1 (left)) which show long-lasting ($\approx 100ms$) activations of pattern and anti-pattern, alternately. The duration of these long plateaus of pattern and anti-pattern activities near the critical τ_1 depends on the value of the parameter β (cf. figure 8). For the smaller values of β (larger noise level) the residence periods are shorter

and for some critical value of β , the saddle-node bifurcation ceases to exist while the intermittency disappears (compare figure 5).

In figure 9 we summarize the network behavior in the parameter (β, τ_{rec}) space. The ferro- and paramagnetic phases denote the regions with stable behavior, i.e. memory and no-memory (noisy) phases, respectively, whereas the oscillatory region represents the novel periodic type of behavior. The dynamics of the synaptic connections, therefore, modifies the phase diagram and produces macroscopic changes of the network behavior.

6 Discussion

We show that incorporation of experimentally observed short-term synaptic dynamics into a stochastic Hopfield network leads to a novel and complex network behavior. Such a network displays bursting activity of memorized patterns and fast switching between memories. Functionality of synaptic depression affects the stability of memory attractors and enables flexible transitions among them. The attractors display a kind of metastability where a small amount of noise can trigger a jump from one state to another. The network dynamics exhibits an oscillatory activity and retrieves patterns or

their mixtures alternately. Although the functional roles and relevance of these results for associative memory are not yet clear, they might be linked to the spatio-temporal encoding scheme (Hoshino et al., 1998) of the olfactory cortex which is believed to represent a good model for analysis of associative memory processes (Haberly and Bower, 1989). In (Hoshino et al., 1998), the model based upon experimental findings showed that the information about odors could be encoded into a spatiotemporal pattern of neural activity in the olfactory bulb, which consists of a temporal sequence of spatial activity patterns. The recognition of olfactory information, thus, as also suggested by other authors, might be related to oscillatory activity patterns. The rapid switching among memory states and the sensitivity to noise resemble the flexibility and metastability of real systems (Kelso, 1995) in receiving and processing new data in continuously changing external conditions. They may also serve as a possible mechanism underlying attention or short-term and working memory (Nakahara and Doya, 1998). In (Nakahara and Doya, 1998) it was emphasized that neural dynamics in working memory for goal-directed behaviors should have the properties of long-term maintenance and quick transition. It was shown that both properties can be achieved when the system parameters are near a saddle-node bifurcation (point) which itself

may be a functional necessity for survival in non-stationary environments.

A similar switching phenomenon has been reported in (Horn and Usher, 1989) where threshold dynamics instead of synaptic dynamics has been considered. Using the threshold function associated with fatigue of a neuron, a self-driven temporal sequence of patterns was obtained. Both adaptive mechanisms, the dynamic threshold and synaptic depression, produce similar effects on the network behavior and enable transitions among stored memories. One may wonder whether these two mechanisms are formally equivalent. In order to explore this question in more details consider the mean field equations $m_i = \tanh(\sum_j w_{ij}x_jm_j + \theta_i)$. The informal equivalence of dynamic synapses and thresholds is that for dynamic thresholds, increasing m_i will decrease θ_i , which decreases future m_i . Similarly for dynamic synapses, increased m_j will decrease x_j , reducing the influence of the product x_jm_j on the network. Therefore, the net effect is similar and indeed dynamic threshold also show oscillations. However, we think that there is no formal equivalence. The reason is that changes in x_j affect the slope of m_j at the right-hand side of the mean field equations, thereby possibly changing the number of fixed points in the system and dynamically changing the system from ferromagnetic to paramagnetic or vice versa. Changes in

θ_i cannot bring about such qualitative changes. These mechanisms as well as some other biological mechanisms (Foss et al., 1997), which also enable noise-induced transitions among attractors, may be related to more dynamic processes of associative memory, the process of associative thinking (Horn and Usher, 1989) or dynamic memory storage (Foss et al., 1997).

Our numerical studies of a network consisting of integrate-and-fire neurons with dynamic synapses show the same switching capability among the memory states as the network with binary neurons. In order to determine the relevance of this behaviour for neurobiology, further analyses of the switching phenomenon and of the phenomenon of intermittency will require more realistic network models as well as more realistic learning rules.

Regarding the learning rule, it has been shown that the covariance learning rule provides optimal memory performance of an associative-memory network with binary neurons (Dayan and Willshaw, 1991; Palm and Sommer, 1996). However, as all other Hebbian forms of plasticity, this rule leads to the excessive growth of synaptic strength. Recent experimental results suggest that there are several regulating mechanisms such as synaptic scaling, spike-timing dependent synaptic plasticity (STDP) or synaptic redistribution (Abbott and Nelson, 2000; Bi and Poo, 1998; Markram and Tsodyks,

1996; Turrigiano et al., 1998) that can prevent uncontrolled synaptic runaway. In (Chechik et al., 2001), it has already been demonstrated that a kind of synaptic scaling, i.e. a neuronal regulating mechanism maintaining the postsynaptic activity near a fixed baseline, improves network memory capacity. The effects of these regulating mechanisms on the network behavior reported here needs to be explored further.

In this study we have focused on a small number of patterns and on the basic mechanism that enables rapid switching among memories. The simulations have shown that noise-induced rapid transitions occur among the (non)overlapping patterns or the mixture states. The transitions become more complex, when the number of patterns is increased and the behavior resembles the motion in a complex multistable system with a large number of co-existing states. An important point is whether this behaviour would be robust to interference among patterns and how it would affect the storage capacity. In (Bibitchkov et al., 2002) it has been shown that the capacity of an infinite network decreases with increasing depression using an equilibrium formulation of the system. It needs to be explored further whether this conclusion is also valid for more realistic neuronal and synaptic dynamics.

Acknowledgments

Dr. J.J. Torres acknowledges support from MCyT and FEDER (“Ramón y Cajal” program) and partial support from Neuroinformatics Thematic Network, Universidad de Granada (“Plan propio”), MCyT (under project BFM2001-2841) and the Dutch Technology Foundation (STW).

7 appendix: Stability matrix

For the system (4.4-4.7) and using the stationary conditions (5.1), the matrix

D becomes:

$$D = \begin{pmatrix} 2m_+(1 - m_+)\beta x_+ & -2m_+(1 - m_+)\beta x_- & 2m_+^2(1 - m_+)\beta & -2m_+(1 - m_+)^2\beta \\ -2m_+(1 - m_+)\beta x_+ & 2m_+(1 - m_+)\beta x_- & -2m_+^2(1 - m_+)\beta & 2m_+(1 - m_+)^2\beta \\ -Ux_+ & 0 & 1 - \frac{1}{\tau_{rec}} - Um_+ & 0 \\ 0 & -Ux_- & 0 & 1 - \frac{1}{\tau_{rec}} - U(1 - m_+) \end{pmatrix} \quad (14)$$

To obtain the eigenvalues of the linearized system we have to solve the characteristic equation

$$|D - \lambda I| = 0, \quad (15)$$

where $|\cdot|$ refers to the determinant and I is the identity matrix. This gives one eigenvalue λ_1 equal to zero. The other three eigenvalues are the solutions

of the cubic equation:

$$\begin{vmatrix} 2m_+(1-m_+)\beta(x_+ + x_-) - \lambda & -2m_+^2(1-m_+)\beta & 2m_+(1-m_+)^2\beta \\ Ux_+ & 1 - \frac{1}{\tau_{rec}} - Um_+ - \lambda & 0 \\ -Ux_- & 0 & 1 - \frac{1}{\tau_{rec}} - U(1-m_+) - \lambda \end{vmatrix} = 0 \quad (16)$$

In general (16) must be solved numerically for an arbitrary solution m_+ . For the particular solution $m_+ = 1/2$ equation (16) reduces to:

$$\left(1 - \frac{1}{\tau_{rec}} - \frac{U}{2} - \lambda\right)^2 \left(\frac{2\beta}{\gamma+2} - \lambda\right) + \frac{\beta U}{\gamma+2} \left(1 - \frac{1}{\tau_{rec}} - \frac{U}{2} - \lambda\right) = 0, \quad (17)$$

which gives

$$\lambda_2 = 1 - \frac{1}{\tau_{rec}} - \frac{U}{2}, \quad (18)$$

and the other two eigenvalues are the solutions of the quadratic equation:

$$\left(1 - \frac{1}{\tau_{rec}} - \frac{U}{2} - \lambda\right) \left(\frac{2\beta}{\gamma+2} - \lambda\right) + \frac{\beta U}{\gamma+2} = 0. \quad (19)$$

References

- Abbott, L. and Nelson, S. (2000). Synaptic plasticity: Taming the beast. *Nat. Neurosci.*, 3, 1178–1183.
- Abbott, L., Varela, J., Sen, K., and Nelson, S. (1997). Synaptic depression and cortical gain control. *Science*, 275, 220–224.
- Amit, D. (1989). *Modeling brain function*. Cambridge University Press, Cambridge.
- Bi, G. Q. and Poo, M. M. (1998). Activity-induced synaptic modifications in hippocampal culture, dependence on spike timing, synaptic strength and cell type. *Journal of Neuroscience*, 18, 10464–10472.
- Bibitchkov, D., Herrmann, J. M., and Geisel, T. (2002). Pattern storage and processing in attractor networks with short-time synaptic dynamics. *Network: Comput. Neural Syst.*, 13, 115–129.
- Bressloff, P. C. (1999). Dynamic synapses, a new concept of neural representation and computation. *Phys. Rev. E*, 60, 2160–2170.
- Chechik, G., Meilijson, I., and Ruppin, E. (2001). Effective neuronal learning with ineffective hebbian learning rules. *Neural Comput.*, 13, 817–840.

- Dayan, P. and Willshaw, D. (1991). Optimizing synaptic learning rules in linear associative memories. *Biol. Cyber.*, 65, 253–265.
- Foss, J., Moss, F., and Milton, J. (1997). Noise, multistability, and delayed recurrent loops. *Phys. Rev. E*, 55, 4536–4543.
- Haberly, L. B. and Bower, J. M. (1989). Olfactory cortex: model circuit for study of associative memory? *Trends Neurosci.*, 12, 258–264.
- Hao, B. L. (1989). *Elementary symbolic dynamics*. (World Scientific Publishing, Singapore.
- Horn, D. and Usher, M. (1989). Neural networks with dynamical thresholds. *Physical Review A*, 40, 1036–1044.
- Hoshino, K., Kashimori, Y., and Kambara, T. (1998). An olfactory recognition model based on spatio-temporal encoding of odor quality in the olfactory bulb. *Biol. Cybern.*, 79, 109–120.
- Kelso, J. S. (1995). *Dynamic patterns: the self-organization of brain and behavior*. MIT Press, Cambridge.
- Kistler, W. M. and Hemmen, J. L. v. (1999). Short-term synaptic plasticity and network behavior. *Neural Comput.*, 11, 1579–1494.

- Liaw, J. S. and Berger, T. W. (1996). Dynamic synapses, a new concept of neural representation and computation. *Hippocampus*, 6, 591–600.
- Markram, H. and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature*, 382, 807–810.
- Miyashita, Y. and Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature*, 331, 68–70.
- Mueller, R. and Herz, V. M. (1999). Content-addressable memory with spiking neurons. *Phys. Rev. E*, 59, 3330–3338.
- Nakahara, H. and Doya, K. (1998). Near saddle-node bifurcation behavior as dynamics in working memory for goal-directed behavior. *Neural Comput.*, 10, 113–132.
- Natschlger, T., Maass, W., and Zador, A. (2001). Efficient temporal processing with biologically realistic dynamic synapses. *Network: Computation in Neural Systems*, 12, 75–87.
- Palm, G. and Sommer, F. (1996). *Models of Neural Networks III*. Springer, Berlin.

- Peretto, P. (1992). *An introduction to the modeling of neural networks*. Cambridge University Press.
- Senn, W., I. S., and Tsodyks, M. (1998). Reading neuronal synchrony with depressing synapse. *Neural Comput.*, 10, 815–819.
- Senn, W., Wyler, K., Streit, J., Larkum, M., Luscher, H. R., Mey, H., Muller, L., Stainhauser, D., Vogt, K., and Wannier, T. (1996). Dynamics of a random neural network with synaptic depression. *Neural Networks*, 9, 575–588.
- Tsodyks, M. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. USA*, 94, 719–723.
- Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.*, 10, 821–835.
- Tsodyks, M., Uziel, A., and Markram, H. (2000). Synchrony generation in recurrent networks with frequency-dependent synapses. *J. Neurosci.*, 20, 50.
- Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C., and Nel-

son, S. B. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, 391, 892–896.

Figure captions

Figure 1: The periodic regimes of the behavior of an associative memory network with dynamic synapses with a single stored pattern ($f = 0.5$). The pattern consists of active ($j = 1, \dots, 60$) and inactive ($j = 61, \dots, 120$) units. Upper panels show a raster plot of the neural activity of all neurons ($N = 120$) in the network during 300 msec. The bottom panels show the overlap m^1 (thin line), the average recovery level of the active units x_+ (solid thick line) and the average recovery level of the inactive units x_- (dashed thick line) as a function of time. For clarity, the dashed thick line is not presented on the right panel.

Figure 2: (Top) Raster plot of a network with $N = 128$ neurons and $P = 6$ orthogonal patterns ($f = 0.5$). The patterns consist of alternating blocks of active units (ones) and inactive units (zeros). The first pattern consists of 2 blocks with size fN , the second pattern consists of 2^2 blocks with size f^2N , the third is composed of 2^3 blocks with size f^3N and so forth. The noise level is fixed to $\beta = 30$ and the time constant for recovery is $\tau_{rec} = 30 \text{ msec}$. Other panels represent overlaps between the current network state and the memorized patterns.

Figure 3: (A) Raster plot of neural activity of a network consisting of $N = 100$ neurons with $P = 10$ non-overlapping patterns during the period of 500 msec. τ_{rec} is set to 75 msec. Each pattern consists of 10 active and 90 inactive units. The first pattern is defined as $\xi_j^1 = 1, \forall j \in (1, \dots, 10)$ and $\xi_j^1 = 0, \forall j \notin (1, \dots, 10)$, the second is given by $\xi_j^2 = 1, \forall j \in (11, \dots, 20)$ and $\xi_j^2 = 0, \forall j \notin (11, \dots, 20)$ and so forth. (B) The panel shows the level of recovery x_j for neuron j ($j = 1, \dots, N$). White regions represent low levels of recovery (depression) while gray (black) parts indicate that synapses are recovering (recovered).

Figure 4: From top to bottom panels represents raster plots of neural activity of the network with $N = 200$ integrate-and-fire units and $P = 5$ patterns (with non-overlapping active bits) for four different values of τ_{rec} and for a fixed noise level of $D = 0.5$.

Figure 5: (Top) The largest eigenvalue $|\lambda|_{max}$ of the linearized (stability) matrix D for a fixed noise level $\beta = 3$ as a function of τ_{rec} (in units of milliseconds). The solid line corresponds to $|\lambda|_{max}$ of the ferromagnetic fixed point $m_+ \neq 0.5$ (memory solutions) and dot-dashed line to $|\lambda|_{max}$ of the paramagnetic fixed point $m_+ = 0.5$ (no-memory solution). τ_1 and τ_2 denote the bifurcation points where $|\lambda|_{max}$ crosses the value 1. (Bottom) The solid line represents three fixed points m_+ as a function of τ_{rec} . For $\tau_{rec} < \tau_1$, the fixed point $m_+ = 0.5$ is an unstable fixed point while $m_+ \neq 0.5$ are stable fixed points. In the $\tau_1 < \tau_{rec} < \tau_2$ region, there exists only one fixed point $m_+ = 0.5$ which is unstable together with a (stable) limit cycle. The white (black) circles correspond to maximal (minimal) amplitudes of oscillations emerging in this region. For $\tau_{rec} > \tau_2$ there is only one stable fixed point solution ($m_+ = 0.5$).

Figure 6: (Top) The largest eigenvalue $|\lambda|_{max}$ of the linearized (stability) matrix D for a fixed level of noise $\beta = 20$ as a function of τ_{rec} . The solid line corresponds to $|\lambda|_{max}$ of the ferromagnetic fixed points $m_+ \neq 0.5$ (memory solutions) and the dot-dashed line to $|\lambda|_{max}$ of the paramagnetic fixed point $m_+ = 0.5$ (no-memory solution). Note that near the first bifurcation point τ_1 , there exist 2 more fixed points. The additional fixed points are unstable and their $|\lambda|_{max}$ -values are represented by the dashed line. At the bifurcation point itself, the stable (solid line) and unstable (dashed line) points coalesce and upon further increase of τ_{rec} disappear entirely. (Bottom) The solid line represents the fixed points m_+ for different values of τ_{rec} . In the $\tau_1 < \tau_{rec} < \tau_2$ region, the circles indicate a presence of an oscillatory phase with white (black) circles corresponding to maximal (minimal) amplitudes of oscillations.

Figure 7: A plot of $m_+(t+1)$ versus $m_+(t)$ corresponding to eq. (8) for $\beta = 20$ and for three different values of τ_{rec} . The intersections of the diagonal and the map (solid line) represent the fixed points (eq. (12)). The (iterative) dynamics of m_+ (black dots) shows that the shape of the limit cycle depends on the value of τ_{rec} . Note that for $\tau_{rec} \gg \tau_1$ the dynamics is uniformly slow and after some transient the black dots form the typical limit cycle (the solid line). For lower values of τ_{rec} , the concentration of the black dots is highest near the places where the ferromagnetic fixed points used to emerge. This means that the dynamics is slower in these areas of the limit cycle and faster in others. The lower right figure represents the enlarged left corner of the lower left figure. Note that the system resides for a long time near the corridors (tangency) between the diagonal and the map when the value τ_{rec} is close to the value τ_1 .

Figure 8: Period of oscillations as a function of the recovery time constant ratio $\frac{\tau_{rec}}{\tau_{rec}^c}$ where τ_{rec}^c denotes the critical value (τ_1) of the recovery time constant (see Fig. 5 and 6). The three different lines correspond to three different noise levels of $\beta = 10, 50$ and 100 .

Figure 9: (Top) Three regimes of behavior in the parameter (β, τ_{rec}) space. τ_{rec} is given in units of milliseconds. The dashed line corresponds to the transition between the memory (ferromagnetic) and oscillatory phase and represents the pairs of critical β and τ_{rec} (τ_1) values. The solid line shows the transition between the oscillatory and noisy (paramagnetic) phase with pairs of critical β and τ_2 values. (Bottom) Inset of lower left part of top figure which shows where the oscillatory phase actually begins. The dot-dashed line is not relevant and represents an approximation artifact since our model (map) is not valid for $\tau_{rec} < 1$.

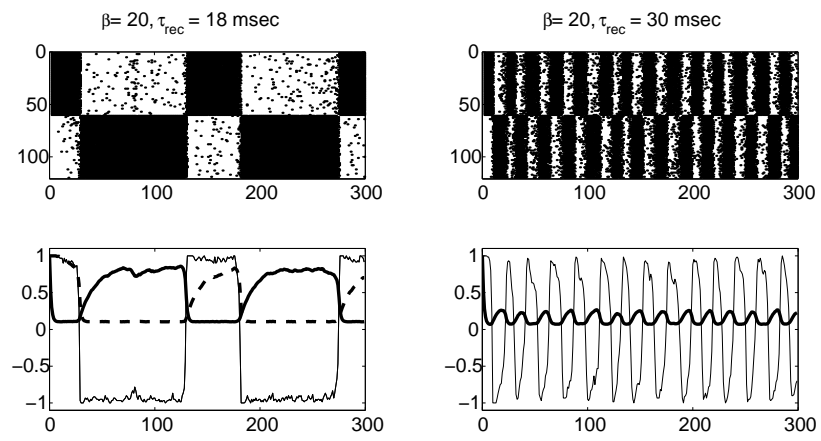


Figure 1, Pantic, NC ms 2448

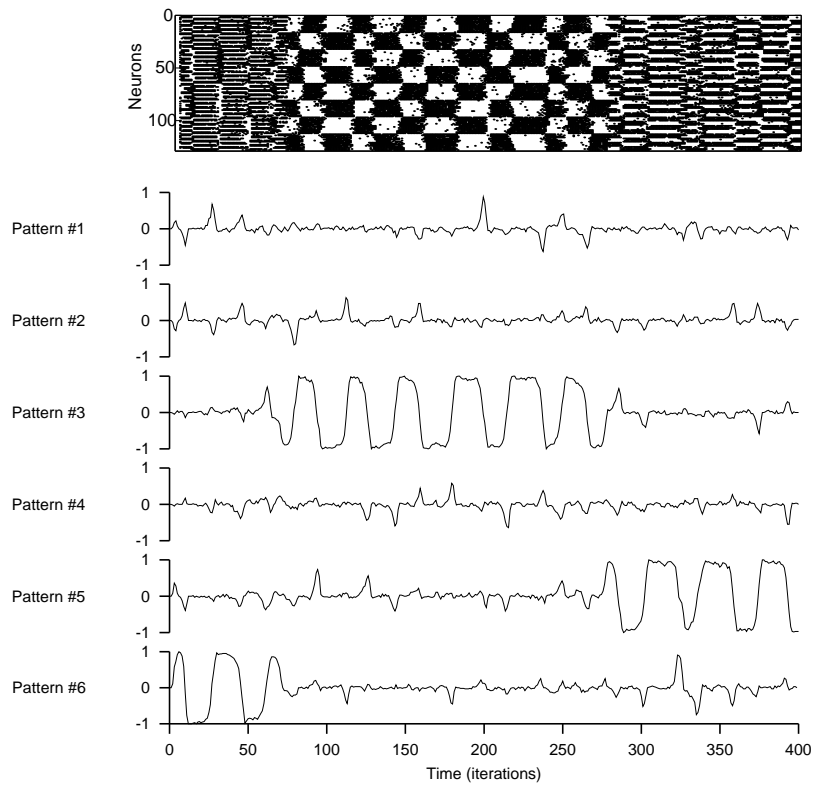


Figure 2, Pantic, NC ms 2448

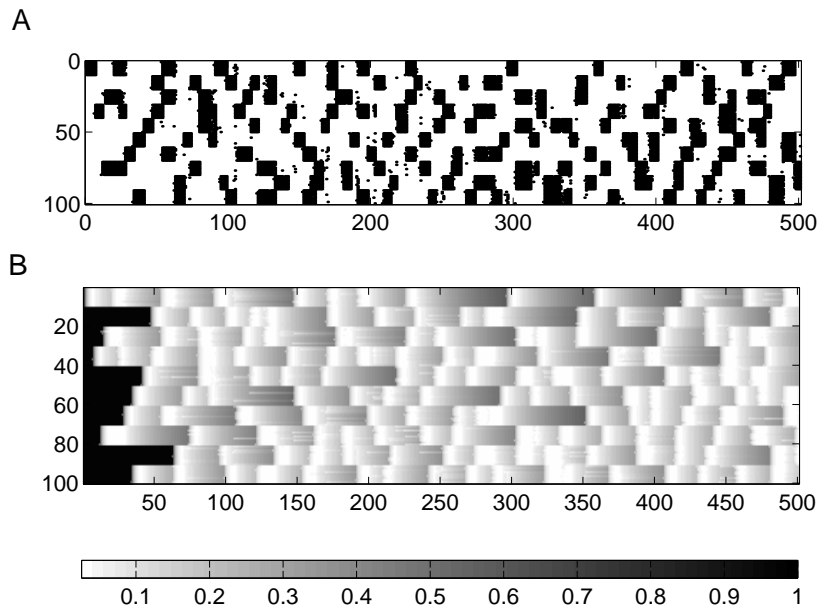


Figure 3, Pantic, NC ms 2448

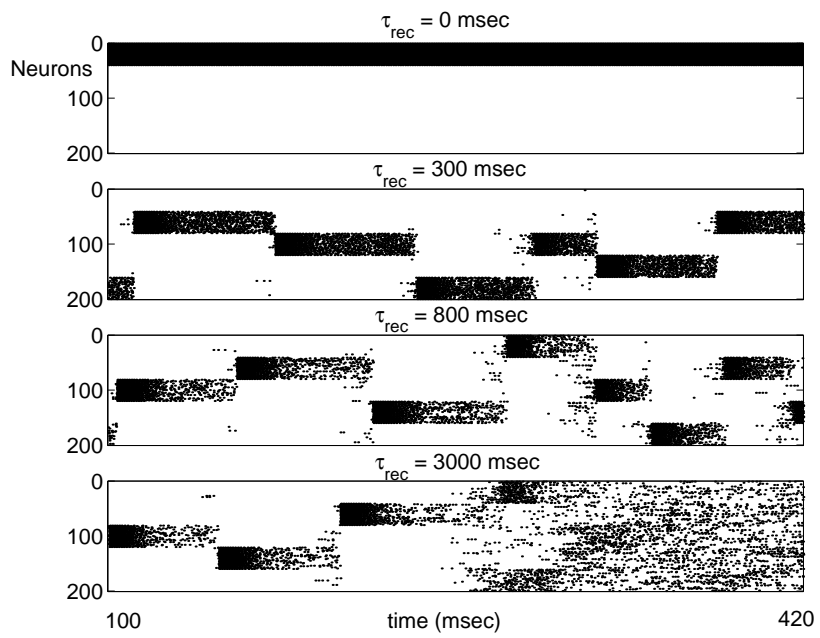


Figure 4, Pantic, NC ms 2448

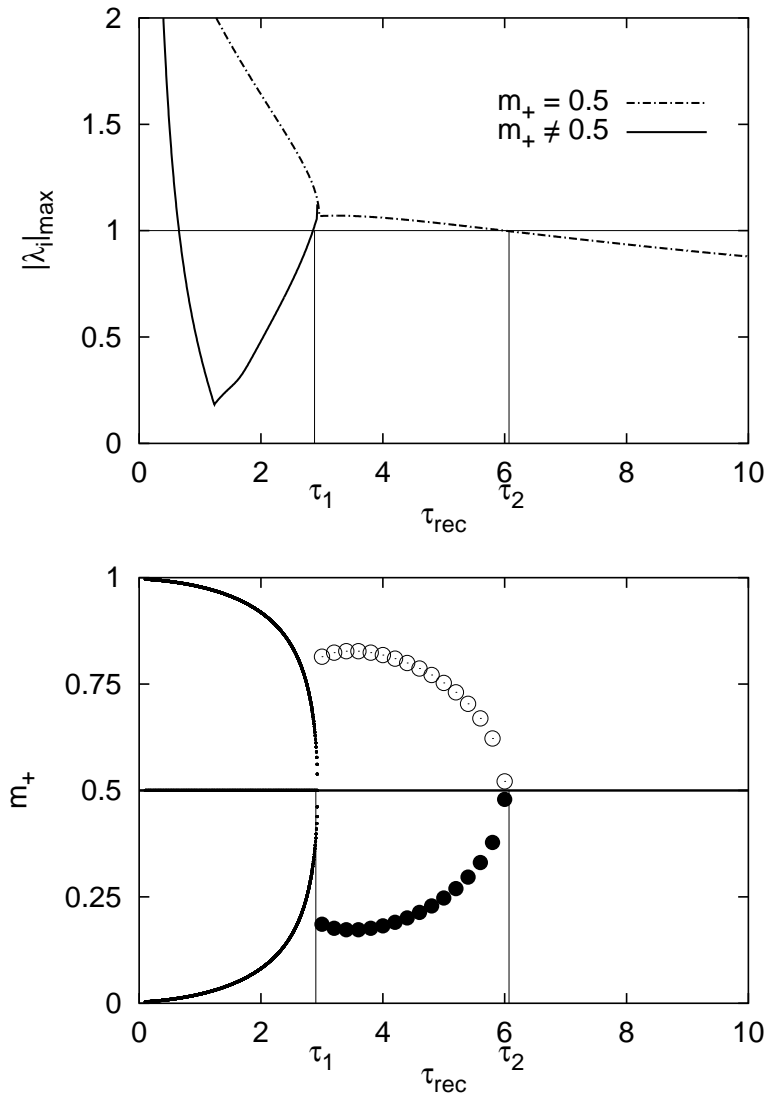


Figure 5, Pantic, NC ms 2448

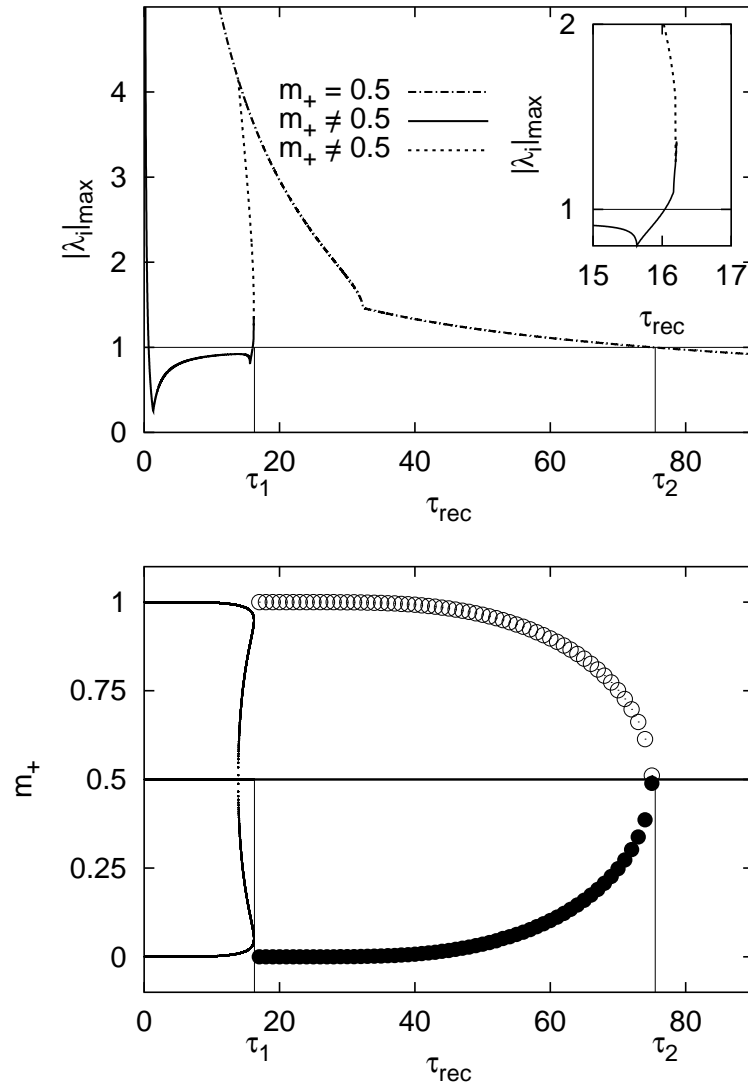


Figure 6, Pantic, NC ms 2448

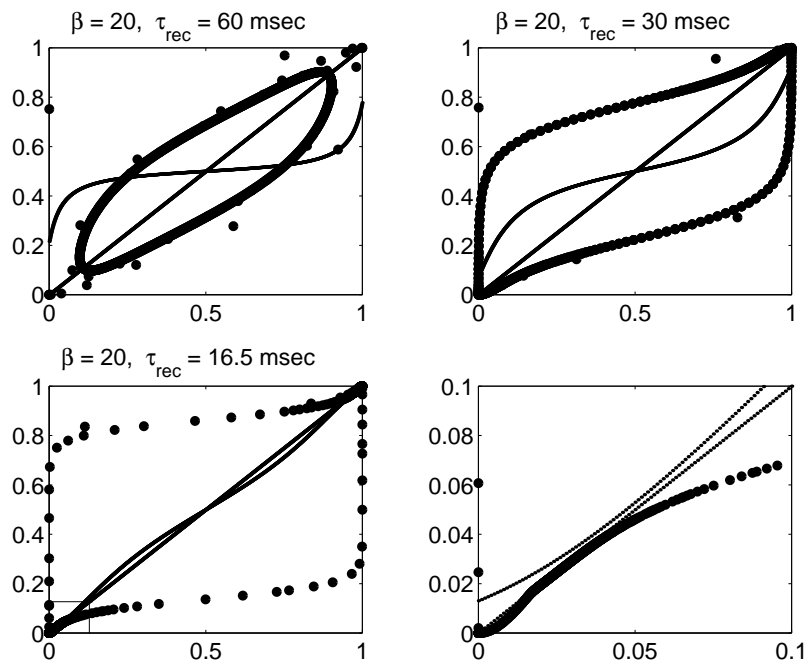


Figure 7, Pantic, NC ms 2448

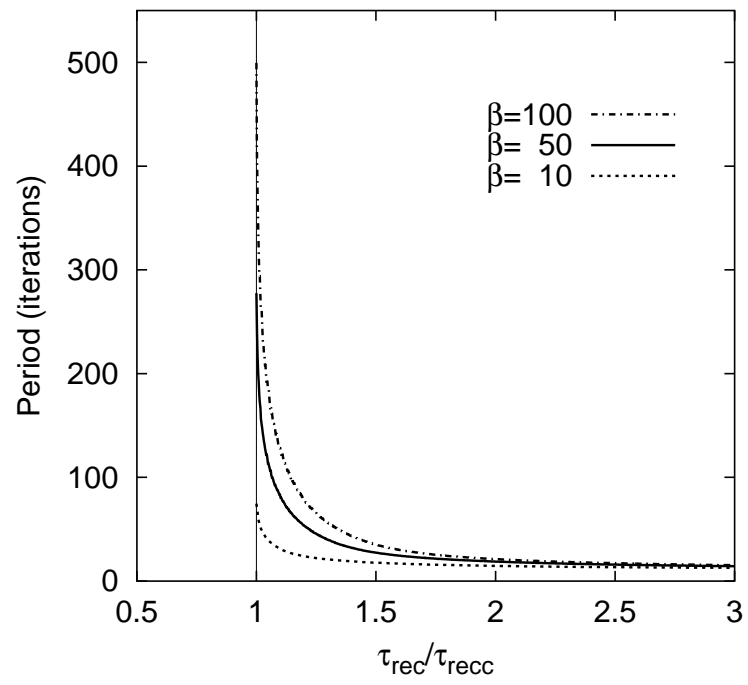


Figure 8, Pantic, NC ms 2448

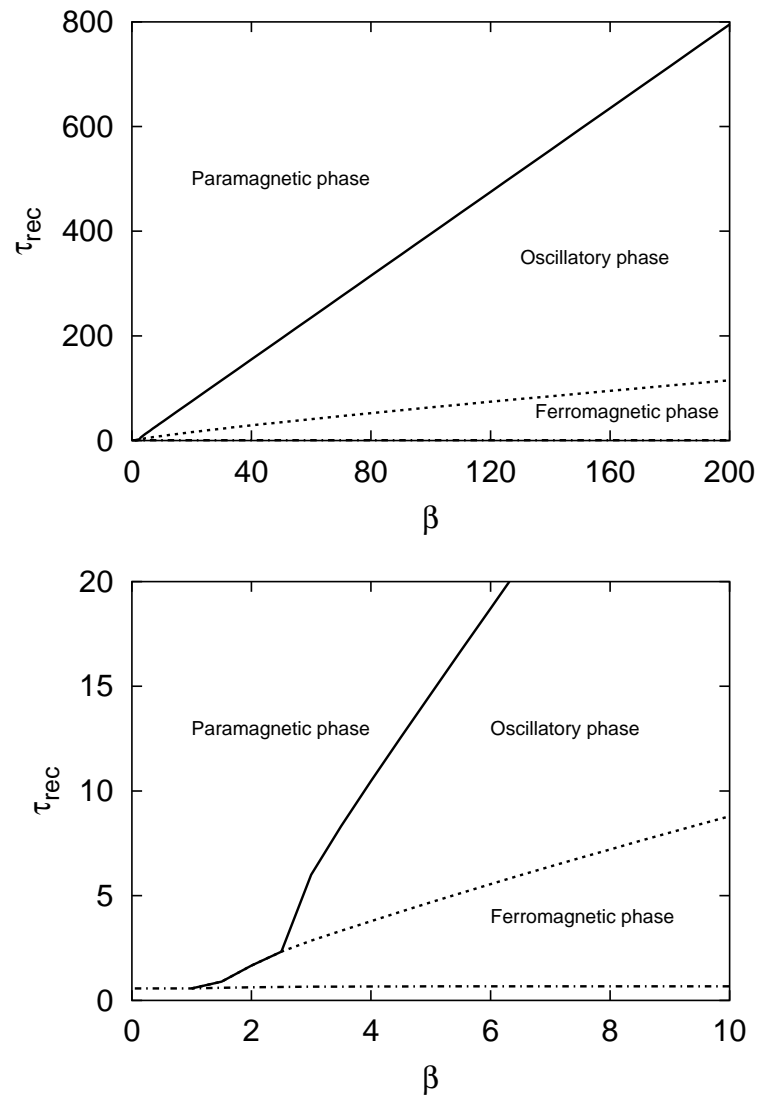


Figure 9, Pantic, NC ms 2448