PERGAMON

Neural Networks 11 (1998) 521–533

Neural Networks

Contributed article

Self-organization and dynamics reduction in recurrent networks: stimulus presentation and learning

Emmanuel Dauce[a], Mathias Quoy[b], Bruno Cessac[c], Bernard Doyon[d], Manuel Samuelides[a,e,*]

[a]*ONERA-CERT/DTIM, 2, avenue Edouard Belin, BP 4025, 31055 Toulouse cedex 4, France*
[b]*ETIS-UCP, 6, av. du Ponceau, 95014 Cergy-Pontoise Cedex, France*
[c]*Institut non linéaire de Nice, 1361 route des lucioles, 06560 Valbonne, France*
[d]*Unité INSERM 455, Service de Neurologie-CHU Purpan, 31059 Toulouse Cedex, France*
[e]*ENSAE, 10, avenue Edouard Belin, BP 4032, 31055 Toulouse cedex 4, France*

Received 13 December 1996; accepted 5 July 1997

Abstract

Freeman's investigations on the olfactory bulb of the rabbit showed that its signal dynamics was chaotic, and that recognition of a learned stimulus is linked to a dimension reduction of the dynamics attractor. In this paper we address the question whether this behavior is specific of this particular architecture, or if it is a general property. We study the dynamics of a non-convergent recurrent model—the random recurrent neural networks. In that model a mean-field theory can be used to analyze the autonomous dynamics. We extend this approach with various observations on significant changes in the dynamical regime when sending static random stimuli. Then we propose a Hebb-like learning rule, viewed as a self-organization dynamical process inducing specific reactivity to one random stimulus. We numerically show the dynamics reduction during learning and recognition processes and analyze it in terms of dynamical repartition of local neural activity.

Keywords: Asymmetric recurrent network; Attractor neural networks; Bifurcations; Chaos; Hebbian learning rule; Non-linear dynamics; Self-organization; Statistical neurodynamics

1. Introduction

Most studies of recurrent neural networks assume sufficient conditions of convergence towards stable state. Convergence is interpreted as a retrieval of a stored pattern (Hopfield, 1982). Models with symmetric synaptic connections exhibit such behavior. Networks with asymmetric synaptic connections lose this convergence property and can have more complex dynamics.

The real brain is a highly dynamical system. Neurophysiological findings have focused attention on the rich temporal structures (oscillations) of neural processes (Eckhorn et al., 1988; Gray et al., 1989) which might play an important role in information processing. Chaotic behavior has been discovered in the nervous system (Babloyantz et al., 1985; Gallez and Babloyantz, 1991). Relying on other neurophysiological results, the study of chaos in neural networks seems very promising in at least two ways: the comprehension of the cognitive processes in the brain (Skarda and Freeman, 1987; Changeux and Dehaene, 1989) and the development of new technologies involving the control of chaos and the massively parallel computability of neural networks.

Freeman's paradigm (Skarda and Freeman, 1987) is that the basic dynamics of a neural system is chaotic and that a particular stimulus is stored as an attractor of a lower dimension than that of the initial chaotic one. The learning procedure thus leads to the arising of such an attractor. During the alert waiting state, the network explores a large region of its phase space through chaotic dynamics. When the learned stimulus is presented, the dynamics are reduced and the systems follows the lower dimensional attractor which has been created during the learning process. The question arises whether this paradigm which has been simulated in Yao and Freeman (1990) using an artificial neural network is due to a very specific architecture or if it is a general phenomenon for recurrent network.

From this point of view, it is of crucial interest to understand the conditions for occurrence of chaos in simple

* Corresponding author. ONERA-CERT/DTIM, 2 Avenue Edouard Belin, BP 4025, 31055 Toulouse cedex, France; e-mail: samuelid@supaero.fr.

0893–6080/98/$19.00
PII: S0893-6080(97)00131-7

recurrent neural networks. A theoretical advance in that direction was achieved by Sompolinsky et al. (1988). They established the occurrence of chaos for fully connected asymmetric random recurrent networks in the thermodynamic limit by using a dynamical mean field theory. In their model, neurons have activation state in $[-1,1]$ with a symmetric transfer function and no threshold. The authors showed that the system exhibits chaotic dynamics. These results were extended by us in Doyon et al. (1993) to the case of diluted networks with discrete time dynamics. We further established (Cessac et al., 1994; Cessac, 1995), in the frame of neural networks with asymmetric couplings and random thresholds, mean-field equations in the approximation of large sized networks. From these equations we inferred different areas with various regimes, which were summarized by a bifurcation map in the control parameters space.

In this paper we focus first on the interaction between random recurrent neural networks and external inputs. We show that when it is submitted to a stimulus, this type of network is driven towards a reduction of its dynamics. The network self-organizes under the constraint of the stimulus and converges to a new attractor which depends on the presented stimulus. If we apply a hebbian learning rule to the network, the stimulus-forced dynamics is reduced to a fixed point within a few steps. The autonomous dynamics remain chaotic and a specific reactivity is observed for the learned stimulus.

The paper is organized in the following manner. In Section 2 we state the main results obtained using mean-field theory in the limit of the large sized network, and we clarify the connections between large size limit and finite size behavior. In Section 3 we investigate the effect of stimulus presentation on the network dynamics, with an overview of some dynamical features which are useful for the understanding of the learning process. In Section 4 we present the results of the learning process and the retrieval properties of the network. We then discuss the results and conclude in Section 5.

2. Random recurrent neural networks

2.1. The random recurrent neural network model

Let us assume we have N neurons in the network. These neurons are connected by asymmetric synapses. J_{ij} is the connection weight of neuron j to neuron i, so $\boldsymbol{J} = (J_{ij})$ is a $N \times N$ matrix. Each neuron computes its *activation state* from its *local field*, which is the sum of the input coming from the other neurons and a threshold.

The dynamics is discrete-time and deterministic. It is governed by the following equations

$$\begin{cases} x_i(t+1) = f[g u_i(t)] \\ u_i(t) = \sum_{j=1}^{N} J_{ij} x_j(t) - \theta_i \end{cases} \tag{1}$$

where $f(x) = (1 + \tanh x/2)$, u_i is the local field and x_i the activation state of neuron i. $x_i(t)$ is a real positive number bounded by 1. In a biological perspective, it may be viewed as the mean firing rate of neuron i at time t. The state of the network at time t is the N-dimensional vector $[x_i(t)]_{i=1\ldots N}$.

So, the *configuration* of the network is defined by the $N^2 + N + 1$ dimensional parameters vector $(\boldsymbol{J},\boldsymbol{\Theta},g)$, where $\boldsymbol{\Theta} = (\theta_i)$ is the N-dimensional vector of the thresholds, and g is the gain of the sigmoid transfer function of the neurons. We are interested in the asymptotics of the non-linear dynamical system Eq. (1). These asymptotics depend on the *configuration parameters vector* $(\boldsymbol{J},\boldsymbol{\Theta},g)$. It is known that this kind of ''Hopfield asymmetric'' network may exhibit chaotic asymptotic behavior. However, systematic predictions of the asymptotics are difficult to assert due to the high dimension of the parameters space even for small sized networks (Renals, 1990). For large sized networks it is impossible to determine the nature of the stationary dynamics for each configuration. Particular models have to be specified which focus the interest on a special region of the parameters space.

One can study models where all the non-vanishing connection weights are sharing one or two typical values (Ginzburg and Sompolinsky, 1994). Another interesting family of models is based on the assumption that the (J_{ij}) and the (θ_i) may be considered as independent sampling of conveniently settled probability distributions (Amari, 1972; Sompolinsky et al., 1988; Cessac et al., 1994). We shall call these networks ''random recurrent neural networks'' (RRNN).

There may be a lot of models belonging to the RRNN family. For instance one can consider randomly diluted networks where the connection weights may vanish with a significant probability (Doyon et al., 1993) or heterogeneous networks with two populations of neurons (Geman, 1982). These models may be chosen that are close to some biological models.

However, here we shall restrict ourselves to a simpler model—homogeneous RRNN. In an homogeneous RRNN, the J_{ij} are a realization of N^2 independent gaussian random variables with expectation $E(J_{ij}) = (\bar{J})/(N)$ and variance $var(J_{ij}) = (J^2)/(N)$. Note that with this scaling law the random variables $W_i = \sum_j J_{ij}$ remain finite when $N \rightarrow +\infty$, with expectation $\bar{J}$ and variance J^2. The thresholds θ_i are a realization of N independent, identically distributed, gaussian, random variables of expectation $E(\theta_i) = \bar{\theta}$ and variance $var(\theta_i) = \sigma_\theta^2$. The 5d vector $(g, \bar{J}, J, \bar{\theta}, \sigma_\theta)$ is called the *distribution parameters vector* of the RRNN.

Practically, if we consider a network of size N defined by the configuration parameters vector $(\boldsymbol{J},\boldsymbol{\Theta},g)$, we can associate a distribution parameters vector $(g, \bar{J}, J, \bar{\theta}, \sigma_\theta)$ by computing the empirical statistics of the configuration parameters vector according to the

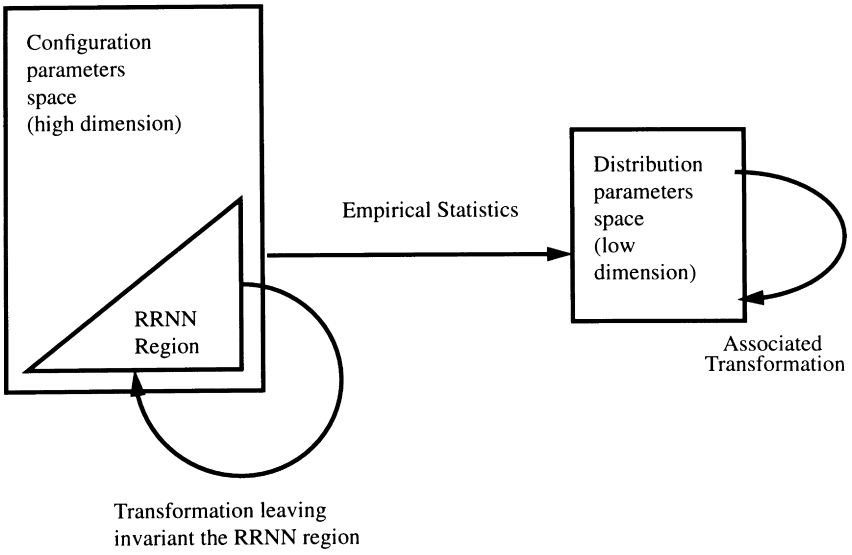

Fig. 1. The passage from the configuration parameters space to the distribution parameters space.

following formula:

$$\begin{cases} [g = g] \\ \bar{J} = \dfrac{1}{N}\sum_{ij} J_{ij} \\ J = \sqrt{\dfrac{N}{N^2-1}\sum_{ij}\left(J_{ij} - \dfrac{\bar{J}}{N}\right)^2} \\ \bar{\theta} = \dfrac{1}{N}\sum_{i} \theta_i \\ \sigma_\theta = \sqrt{\dfrac{1}{N}\sum_{i}(\theta_i - \bar{\theta})^2} \end{cases}$$

Some simple transformations on the configuration parameters space such as the constant dilation of the connection weights ($J_{ij} \rightarrow \alpha J_{ij}$), the constant dilation of the thresholds ($\theta_i \rightarrow \alpha\theta_i$) and the constant shift of the thresholds ($\theta_i \rightarrow \theta_i + \tau$) induce associate transformations on the distribution parameters space. For instance, the transformation ($J_{ij} \rightarrow \alpha J_{ij}$) of the configuration parameters space induces the following transformation of the distribution parameters $(g, \bar{J}, J, \bar{\theta}, \sigma_\theta) \rightarrow (g, \alpha\bar{J}, \alpha J, \bar{\theta}, \sigma_\theta)$ and so on.

In the configuration parameters space, we can define the RRNN region as the fuzzy subset of the ($\boldsymbol{J}$,$\boldsymbol{\Theta}$,g) such that the J_{ij} and the θ_i satisfy the requirements for independent sampling of respective normal laws $N(\bar{J}, J)$ and $N(\bar{\theta}, \sigma_\theta)$ with respect to usual statistical tests of given levels of confidence. Indeed these levels determine the fuzziness degree of belonging to the RRNN region. The previous transformations *are leaving the RRNN region invariant*. Fig. 1 summarizes this point of view.

Actually, the transformation $(g, \bar{J}, J, \bar{\theta}, \sigma_\theta) \rightarrow (gJ, (\bar{J})/(J), 1, (\bar{\theta})/(J), (\sigma_\theta)/(J))$ left the evolution equation unchanged, so the true dimension of the distribution parameters space of homogeneous RRNN is four. In this paper we shall focus on the specific case where $\bar{J} = 0$ and $\bar{\theta} = 0$. The dimension of the parameters space is then reduced to two. Setting $J = 1$, we shall essentially study the effect of the relative size of the neural gain g and the threshold standard deviation σ_θ.

We shall show hereafter that Freeman's paradigm about stimulus-forced self-organization by reduction of dynamics, that was demonstrated on a specific functional architecture for the olfactive bulb, is valid in this simple and general frame model. The methods which are used here can be extended to more structured architecture of RRNN. This issue will be addressed in the final discussion.

2.2. The mean-field theory for RRNN

The dynamics of homogeneous RRNN can be studied analytically in the large size limit (i.e. when the size N of the system goes to infinity). This model obeys a *propagation of chaos* principle (Cessac et al., 1994). Namely, the intra-correlation between neural activities within finite subsets of the network is vanishing. So one can infer that the neural local fields $t \rightarrow u(t)$ are independent, identically distributed gaussian processes in the large size limit. This property was introduced by Amari (1972) in the field of neural networks. More recently, rigorous results of propagation of chaos for high dimensional random dynamical systems have been obtained by Sznitzman (1984); Ben Arous and Guionnet (1995). It is the main analytical tool in the study of RRNN.

The equations determining the gaussian process of neural activity are called the mean-field equations. Mean-field equations for various models of RRNN were studied by several authors (Sompolinsky et al., 1988; Molgedey et al., 1992). In previous papers (Cessac et al., 1994; Cessac, 1995) we extended the works of these authors and computed the characteristics of the dynamics of our model. The results are summarized hereafter.

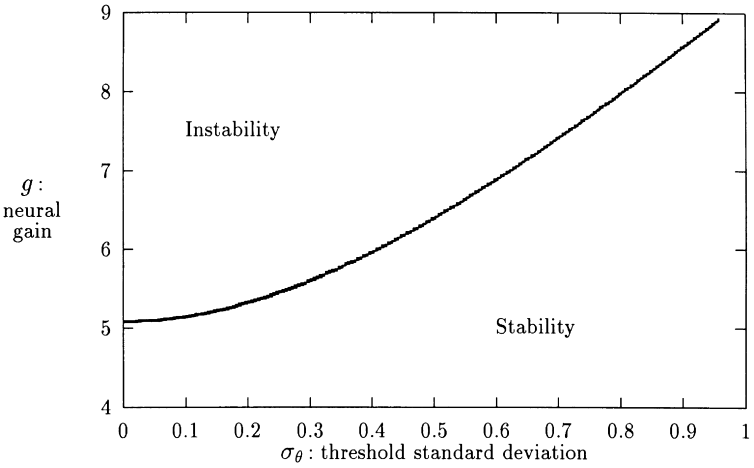

Fig. 2. The theoretical stability map of the RRNN model. The two distribution parameters, σ_θ (threshold standard deviation) and g (neural gain), are reported on the coordinate axis.

The moments of the neural local fields (u_i) and of the neural output (x_i) are obviously equidistributed if the initial condition respects this random symmetry.

$$\text{Let}\begin{cases}\mu(t)=E[u(t)], v(t)=\text{var}[u(t)], \Delta(t,t')=\text{cov}[u(t),u(t')]\\ m(t)=E[x(t)], q(t)=E[x(t)^2], C(t,t')=E[x(t),x(t')]\end{cases}$$

The average is taken over the randomness of connection weights and thresholds. Note that the set average is well approximated by the space average obtained by taking the mean values of the individual values over large subsets of neurons thanks to the propagation of chaos property. Evolution equations have been derived for these moments. These equations define a dynamical system—the mean-field dynamical system.

From now on we shall restrict ourselves to $\bar{J}=0$, $J=1$ and $\bar{\theta}=0$. Then we have two independent parameters that determine the statistics of the network for large sized networks, namely (σ_θ,g). First let us consider the evolution equations of [$\mu(t)$, $v(t)$, $m(t)$, $q(t)$]. For each value of the parameters there is one and only one fixed point for that recurrence (Cessac et al., 1994). Let it be (μ_∞, v_∞, m_∞, q_∞) at this fixed point.

If there is a stable fixed point $\boldsymbol{\chi}*=(x_i^*)$ for evolution Eq. (1) of large size RRNN, then according to propagation of chaos principle, this fixed point has empirical statistics close from (μ_∞, v_∞, m_∞, q_∞) More precisely, if we set $u_i^*=(1)/(g)f^{-1}(x_i^*)$, the (u_i^*) are a realization of independent gaussian variables of mean μ_∞ and variance v_∞. So it is possible to perform the stability analysis in the large-size limit. Indeed, the Jacobi matrix of the evolution operator of Eq. (1) is

$$\boldsymbol{D}(\boldsymbol{\chi}*)=2g\boldsymbol{L}(\boldsymbol{\chi}*)\boldsymbol{J}$$

where $\boldsymbol{L}(\boldsymbol{\chi}^*)$ is the diagonal matrix of components $(x_i^*-x_i^{*2})$.

The spectral radius of this random matrix $\rho[\boldsymbol{D}(\boldsymbol{\chi}^*)]$ can be computed (Cessac et al., 1994). The equation $\rho[\boldsymbol{D}(\boldsymbol{\chi}^*)]=1$ determines a critical line in the 2d parameters space. This line is represented in Fig. 2. It is the boundary between the unstable dynamics domain in the upper part of the plane (large g, low σ_θ) and the stable dynamics domain in the other part (low g, large σ_θ).

When $\rho[\boldsymbol{D}(\boldsymbol{\chi}^*)]>1$, the mean-field theory allows us to determine the state dynamics in the large-size limit. In (Cessac, 1995), the asymptotic evolution of $\Delta(t,t')$ has been studied and for $t\neq t'$ was shown to converge towards a constant Δ_∞. That means that the stationary dynamics of the (u_i) are independent identically distributed gaussian processes defined by:

$$u_{i,\infty}(t)=u_i^*+b_i(t)$$

where the b_i are independent, identically distributed, centered white noises of variance $v_\infty-\Delta_\infty$ and where the (u_i^*) *are independent variables distributed according to a normal law* $\mathbf{N}(\mu_\infty, \Delta_\infty)$. The interpretation of this result in terms of dynamical systems is straightforward.

Let us study the mean quadratic distance between two trajectories $\boldsymbol{u}^{(1)}(t)=(u_i^{(1)}(t))$ and $\boldsymbol{u}^{(2)}(t)=(u_i^{(2)}(t))$ originating from different initial conditions.

$$E[\|\boldsymbol{u}^{(1)}(t)-\boldsymbol{u}^{(2)}(t)\|^2]=\frac{1}{N}\sum_{i=1}^{N}E\{[u_i^{(1)}(t)-u_i^{(2)}(t)]^2\}$$

$$=2\{var\ u_i(t)-cov[u_i^{(1)}(t),u_i^{(2)}(t)]\}$$

If we set $\Delta^{1,2}(t)=\text{cov}\,[\boldsymbol{u}^{(1)}(t),\boldsymbol{u}^{(2)}(t)]$ it has been shown in (Cessac, 1995) that the evolution equation that governs the $\Delta^{1,2}$ is identical to the equation that governs Δ and thus $\lim_{t\to\infty}\Delta^{1,2}(t)=\Delta_\infty$. Therefore, in the stationary regime the mean quadratic distance between two trajectories $\boldsymbol{u}^{(1)}(t)$ and $\boldsymbol{u}^{(2)}(t)$ is converging towards a limit value, $2(v_\infty-\Delta_\infty)$, that does not depend on the distance between the initial conditions. So, in that sense, the model is sensitively dependent on initial conditions. Trajectories originating from close initial conditions are diverging. This method was

introduced by Derrida and Pomeau to study another kind of recurrent random network, the Kauffman model (Derrida, 1988).

2.3. The finite size effects

In our simulations, the predictions of the theory about the qualitative behavior of the network in the large size limit are checked with a good accuracy as soon as $N > 50$. In (Cessac et al., 1994) quantitative predictions given for m_∞ by mean-field theory are compared with the measurement of $m_{\text{net}}(t) = (1)/(N)\sum_{i=1}^{N} x_i(t)$ for networks of 100 neurons.

There is very good agreement between simulated data and theoretical results. For instance, if we consider Fig. 2 we can predict that in the neighborhood of the reciprocal image of the bifurcation line of the distribution parameters space in the RRNN region of the configuration parameters space (see Section 2.1) the asymptotic dynamics of the networks undergoes qualitative change. This is actually the case.

However, there is a crucial difference between MFT (mean-field theory) predictions and the results of simulations. MFT establish a sharp transition between a fixed point dynamics and a stochastic regime. Finite size systems exhibit a wider variety of stationary dynamics. *Simulations show the evidence of a quasi-periodicity route to deterministic chaos*. For instance, in Eq. (1), when g is increased and the other parameters are kept constant, a fixed point is destabilized by a Hopf bifurcation and gives rise to a limit cycle. Then another Hopf bifurcation occurs leading to a 2-torus densely covered by the trajectories. Frequency locking follows leading to chaos. Eventually, the dimension of the strange attractor of the chaotic regime increases up to values where its computation within reasonable computational time is not possible any more.

When the number of neurons increases, this transition zone becomes smaller and smaller and one approaches the sharp transition to stochasticity in the large size limit. For instance, we compute for $J = 1$, $\bar{\theta} = 0$ and $\sigma_\theta = 0$ the theoretical value of g for the bifurcation by using MFT equations. The result is $g = 5.08$. We compared the result with the transition interval from stable fixed point to chaos (characterized by the sensitivity to initial conditions along the trajectory) obtained by numerical simulation of networks dynamics for increasing N. We obtain the following results, shown in Table 1, averaging the simulation results over 30 networks in each case.

Table 1
Transition intervals: the lower bound is the value of g for which the fixed point becomes unstable and the upper bound is the value of g for which a strange attractor appears

Size of the network	Transition interval
128	[4.79, 5.85]
256	[4.81, 5.78]
512	[4.72, 5.30]

Therefore, even if the MFT does not give the detailed structure of this cascade of bifurcations, the range of parameters corresponding to this increase of complexity in the dynamics can be statistically located with an increasing accuracy with increasing size N. So the MFT bifurcation map helps us to classify the dynamics of the networks in the RRNN region.

In Section 3 we shall study the effect of a random stimulus presentation to the network. Since this transformation leaves the RRNN region invariant, its effect can be predicted by MFT.

In Section 4 we shall study the effect of a hebbian learning rule. The hebbian learning rule breaks the independence of the connection weights and the network configuration is driven by the learning process outside of the RRNN region quite quickly. In this case the MFT theory cannot be used to predict the behavior of the network.

3. Stimulus presentation and dynamics of RRNN

In this section we shall investigate the effect of stimulus presentations to recurrent networks. From now on we restrict the range of the size of these networks from 100 up to 1000 neurons. Within these limits, the quantitative predictions of MFT are checked quite accurately. The transition zone between fixed point and stochasticity, including the presence of limit cycles and strange attractors of small fractal dimension, is still large enough to keep the network configuration within this zone while a stimulus presentation process or a learning process is performed.

3.1. Stimulus-forced reduction of dynamics

We now study the reaction of the network when it is submitted to an external input—*the stimulus*, which is imposed to the network while the dynamical process is going on. It is not a ''flash'', leading to some kind of transient relaxation phenomenon. In the field of attractor neural networks information processing, this strategy seems more realistic than other models (like the Hopfield one) for which the presentation is only a change in the initial conditions of the neural states. The presentation of a stimulus is equivalent to modifying the dynamical system itself. So each stimulus presentation creates a new system with a different attractor. The attractor (final state of the network) will be an output of the confrontation of the stimulus and the processing system. Let us call *autonomous dynamics* the network dynamics without any external input and *stimulus-forced dynamics* the dynamics of the new system which is obtained when the network is submitted to an external input.

The stimulus is a vector (a pattern) of random gaussian variables I_i, independent of the θ_is, with a mean value $\bar{I}$ and a variance σ_I^2. Since *there is no geometry* in our neural network (homogeneous RRNN), the statistics of the connection matrix is invariant by any permutation of the neurons. The

nature of the chosen patterns (random independent sampling of a normal law) is coherent with the homogeneous RRNN structure of the network. The only relevant stimulus characteristic is the empirical distribution of the pattern value. When a stimulus is presented, the evolution of the network is now governed by the following evolution equation:

$$\begin{cases} x_i(t+1) = f[gu_i(t)] \\ u_i(t) = \sum_{j=1}^{N} J_{ij}x_j(t) - \theta_i + I_i \end{cases} \quad (2)$$

Most of the time the presentation of a stimulus induces a reduction in dynamics. This phenomenon was independently observed by us (Quoy et al., 1993) and by Palmadesso and Dayhoff (1995). It was called ''attractor locking'' by these last authors in the case where the stimulus evoked attractor is a limit-cycle. Fig. 3 displays an example of this *stimulus-forced reduction in dynamics.* Hereafter, this property is systematically investigated.

We start our first simulation with a network whose

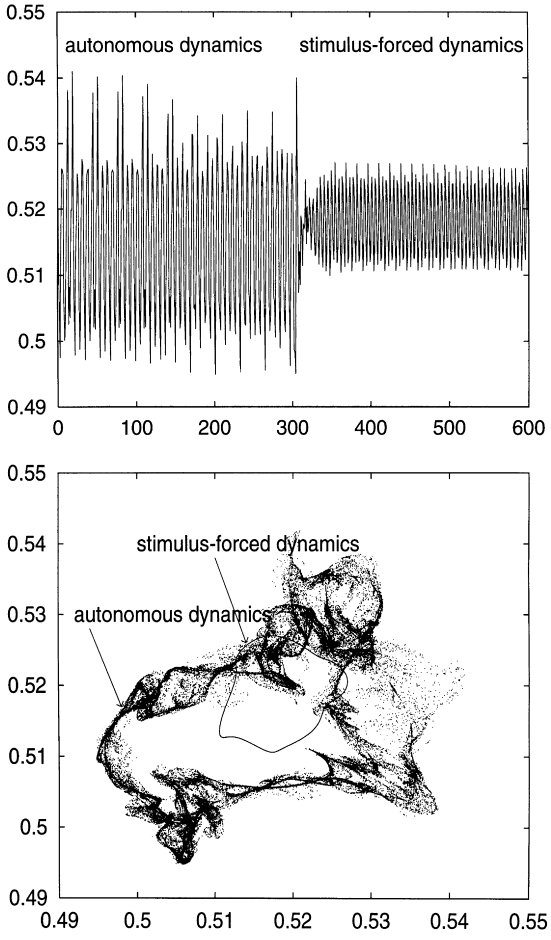

Fig. 3. Example of stimulus-forced reduction of dynamics. (A) Time variation, after 500 transients, of the observable $m_{\text{net}}(t)$. The stimulus is presented at time $t = 300$. (B) Representation of the two attractors corresponding to the autonomous dynamics and the stimulus-forced dynamics in the $[m_{\text{net}}(t), m_{\text{net}}(t+1)]$ phase plane (plotted on 100 000 points).

Table 2
Critical values for destabilization and arising of chaos on RRNN. Simulation results for $N = 200$ averaged over 50 networks

σ_I	g_{dest} empirical mean	g_{dest} empirical s.d.	g_{chaos} empirical mean	g_{chaos} empirical s.d.	g MF T
0	5.15	0.96	5.84	0.92	5.08
0.2	5.30	1.35	6.46	1.76	5.32
0.4	6.04	1.35	7.11	1.45	5.96
0.6	7.58	2.57	8.61	2.81	6.88
0.8	8.81	3.14	10.41	3.26	7.97
1	10.56	4.65	12.17	4.39	9.18

distribution parameters are $g = 6.2$, $\bar{J} = 0$, $J = 1$, $\bar{\theta} = 0$, $\sigma_\theta = 0$. The gain g has been chosen to be large enough for the dynamics to be chaotic, close from the edge of chaos. The size of the network is $N = 200$. The stimulus is generated with $\bar{I} = 0$, $\sigma_1 = 0.7$. The chosen observable is the time-dependent space average of the neural activity: $m_{net}(t) = (1)/(N)\sum_{i=1}^{N} x_i(t)$. Fig. 3A. represents the plot of $m_{\text{net}}(t)$ versus t before and after the presentation of the stimulus. Fig. 3B represents the plots of $[m_{\text{net}}(t), m_{\text{net}}(t+1)]$ in the stationary regimes associated to the autonomous dynamics and the stimulus-forced dynamics (two-dimensional projection).

This simulation allows us to point out a general phenomenon: the dynamics tends to be more complex with a growing g and to reduce with a growing σ_I. To check this assertion we select randomly 50 networks of size 200. Then we run the networks for various values of the parameters g and σ_I, and determine the nature of the asymptotic dynamics. More precisely, for each of the following values for σ_I: 0, 0.2, 0.4, 0.6, 0.8 and 1 we determine the critical values g_{dest} and g_{chaos} for destabilization of the fixed point and the occurrence of chaos (characterized by the sensitivity to initial conditions). These values are averaged on 50 random configurations for each value of σ_I. The results are displayed in Table 2.

We observe a general confirmation of the stimulus-forced reduction of dynamics. However, we still find relatively large individual variations among the samples. We also observe a tendency to the increase of these variations for large at σ_I.

This phenomenon can be predicted by the MFT bifurcation map (see Fig. 2) of the system in the distribution parameters space. Indeed, adding one input is equivalent to changing the statistics of the threshold. The global resulting threshold has a mean value of $\bar{\theta} - \bar{I}$ (0 in our simulations) and a variance of $\sigma_\theta^2 + \sigma_I^2$. Its statistical properties (independent sampling of a normal law) are unchanged. We can therefore interpret the presentation of a stimulus as a transformation acting on the configuration parameters space, leaving the RRNN region invariant and inducing a simple translation on the distribution parameter's space. Depending on the distribution parameters, the network configuration is more or less near the boundary of the chaotic region. The

presentation of a stimulus is equivalent to increasing σ_θ. So if the network configuration is near this boundary (the edge of chaos), the presentation of the stimulus will tend to reduce the dynamics of the system by crossing the boundary and falling into the $T2$ torus and limit cycle area.

3.2. Stimulus-forced self-organization

The stimulus-forced reduction of dynamics can be explained in the following way. The MFT bifurcation map shows that dynamics are reduced when g is decreasing or when σ_I is increasing. The causes of these dynamics reductions are not the same. When g is small, most of the neuron transfer functions are working in the linear regime and the fluctuations of the set of neurons are reduced by the low gain. In contrast, the stimulus-forced reduction of dynamics is a non-linear effect.

Actually, when the absolute value of the threshold (or stimulus) on one neuron is large, its mean local field u_i^* tends to be out of the linear range of its transfer function. It remains either strongly negative or strongly positive. Such a neuron is either *silent* $[u_i^* \ll 0 \Rightarrow x_i(t) \approx 0, \forall t]$ if its mean local field is strongly negative or *saturated* $[u_i^* \gg 0 \Rightarrow x_i(t) \approx 1, \forall t]$ if its mean local field is strongly positive. While reporting our simulations, we shall call a neuron such that the activation state $x_i(t) < 0.01, \forall t$, silent, and a neuron such that $x_i(t) > 0.99, \forall t$, saturated. If a neuron remains silent or saturated, its activation state does not vary any more. A neuron which is neither silent nor saturated is called *dynamical*. We shall show that the stimulus-forced reduction of dynamics is caused by the reduction of the proportion of dynamical neurons.

Let us first observe the influence of the variation of one particular neuron stimulus by acting on the individual stimulus of that neuron. We build a network of $N = 200$ neurons with $g = 6$, $\bar{J} = 0$, $J = 1$, $\bar{\theta} = 0$, $\sigma_\theta = 0$ which exhibits a chaotic behavior, and we select one neuron i. Then we vary its individual stimulus I_i and compute its time average activation state $\langle x_i(t) \rangle$ and standard deviation $\langle x_i(t)^2 - \langle x_i(t) \rangle^2 \rangle^{1/2}$ versus I_i.

In Fig. 4 we observe that:

- for $0.8 < I_i$, the neuron i is saturated;
- for $-0.2 < I_i < 0.8$, the neuron i is dynamical (i.e. fluctuating with time);
- for $I_i < -0.2$ the neuron i is silent.

Note that the range of values of dynamical behavior is not centered around 0. This can be interpreted by the inner input of the other neurons. Actually some neurons of the network are silent. Others are saturated and their contribution to the dynamics of other neurons through the connection weights *is equivalent to an additional threshold*. This additional threshold is variable for each neuron. It is roughly equal to -0.3 for the considered neuron on Fig. 4. The *effective threshold* of a neuron is the sum of the initial threshold, minus the stimulus local intensity, and *the input coming from saturated neurons* through connection weights.

Moreover, if σ_I is increasing (the ''intensity'' of the stimulus is increasing), the dynamics reduction is not a simple mechanical consequence of the threshold shift as it

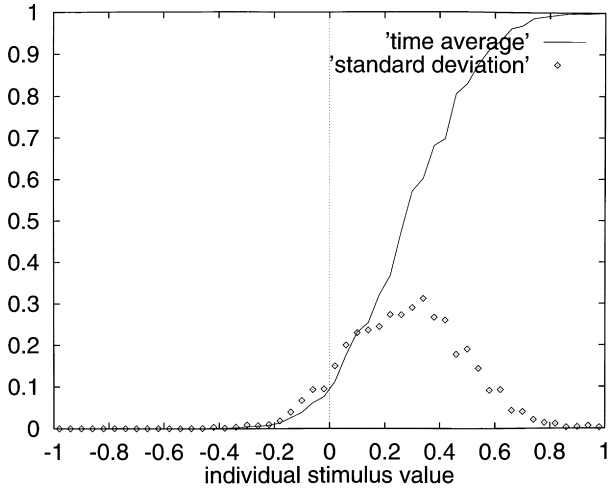

Fig. 4. Variation of the time average activation state and standard deviation of one individual neuron dynamics, versus its individual stimulus value.

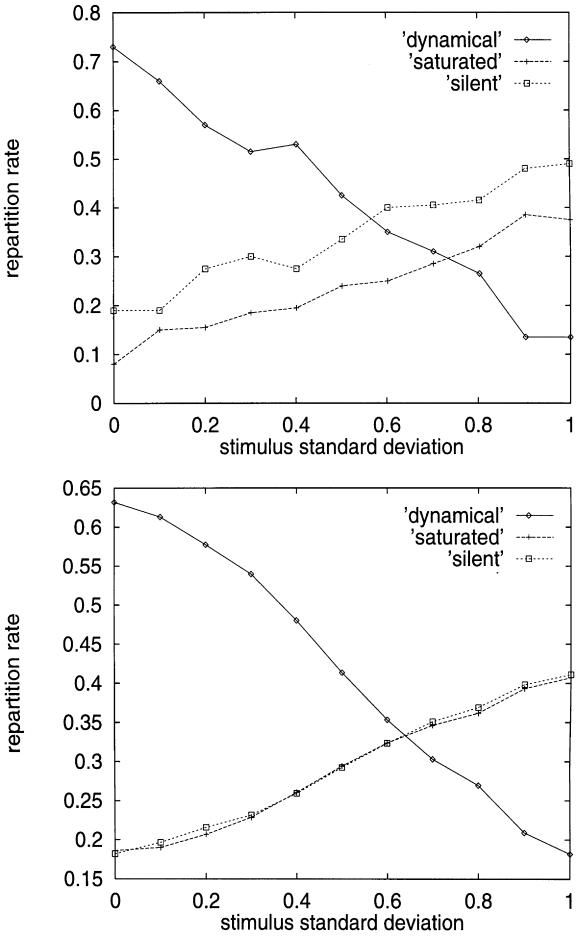

Fig. 5. Variation of the dynamical repartition of neurons when the intensity of the stimulus is increased. (A) Repartition for a typical RNNN. (B) Repartition averaged over 50 networks.

would be for a single neuron. It is a collective reinforcement effect. When a neuron becomes saturated due to a stimulus individual high intensity, its contribution to the input of the other neurons is acting exactly like a threshold through the connection weights. Then the intensity of the effective threshold is enhanced for every neuron and more and more neurons are driven to a silent or saturated state (Fig. 5).

4. Hebbian learning and self-organization of RRNN

In this paragraph we show that the application of a hebbian learning rule to any RRNN leads quickly to a drastic reduction of dynamics. Though this result is general, the speed of this process is variable due to the specificity of each association network-stimulus. Many neurophysiological studies stress the role of synaptic plasticity in the learning process. Two main schemes are proposed to explain the synoptic efficacy modification—the correlation between the activities of pre- and post-synaptic neurons [and its common translation as the Hebb rule (Hebb, 1949)], and the conjunction of arrival of spikes on a neuron (Alkon et al., 1990). As we only consider mean firing rates, we will only work with the first scheme, and propose Hebb-like learning rules.

Since the learning rule consists in correlating the connection weights, MFT theory is no more valid in the large size limit. So all the assertions of this paragraph are based on simulations. Then we show that this reduction of dynamics is specific to the presented stimulus. In that sense, the network has learned the stimulus. However, this effect is subject to large quantitative variations.

4.1. Hebbian self-organization and dynamics reduction

We selected the following hebbian learning rule:

$$\text{if } x_j(t) > 0.5, \text{ then } \Delta J_{ij} = \frac{\alpha}{N}[x_j(t) - 0.5][x_i(t+1) - 0.5]$$

$$\text{else } \Delta J_{ij} = 0$$

where α is a small positive real number called the learning rate parameter.

There are two time scales during the learning process: a fast one for the iteration of the network dynamics and a slow one for the iteration of learning. In our simulations we have one learning step every 100 dynamics steps.

We add the constraint that a weight can not modify its sign. At each learning step, J_{ij} is replaced by $J_{ij} + \Delta J_{ij}$ if the signs of these two terms are identical.

This learning rule means that when the sending neuron j is active (i.e. its activation state is superior to 0.5) and the receiving neuron i is active too, then the weight *increases*. For excitatory synapses ($J_{ij} > 0$), this rule leads to an increase of the absolute value of the weight, and for inhibitory synapses ($J_{ij} < 0$) it leads to a decrease of the absolute value of the weight. If the sending neuron j is active and the receiving neuron i is inactive (i.e. its activation state is inferior to 0.5) then the weight *decreases*. If the sending neuron is inactive, the weight remains unchanged.

- $x_i(t+1) > 0.5 \Rightarrow \Delta J_{ij} \geqslant 0, \forall j$
- $x_i(t+1) < 0.5 \Rightarrow \Delta J_{ij} \leqslant 0, \forall_j$

Of course one can imagine learning rules with different thresholds instead of 0.5, or a smooth version of previous rule replacing the ''IF'' condition by a sigmoidal function of the activation of the sending neuron. It is also possible to change the time-scale of the learning process. In that case, we average the activation state of the neuron over the attractor. The basic results of the paragraph are not changed by these variants.

In our simulations, we have generated RRNNs according to the following specifications: $N = 200$ neurons, $g = 15$, $\bar{J} = 0$, $J = 1$, $\bar{\theta} = 0$ and $\sigma_\theta = 0$. For these values of the distribution parameters, the gain value is large enough to ensure that the autonomous dynamics is always chaotic (see Table 2).

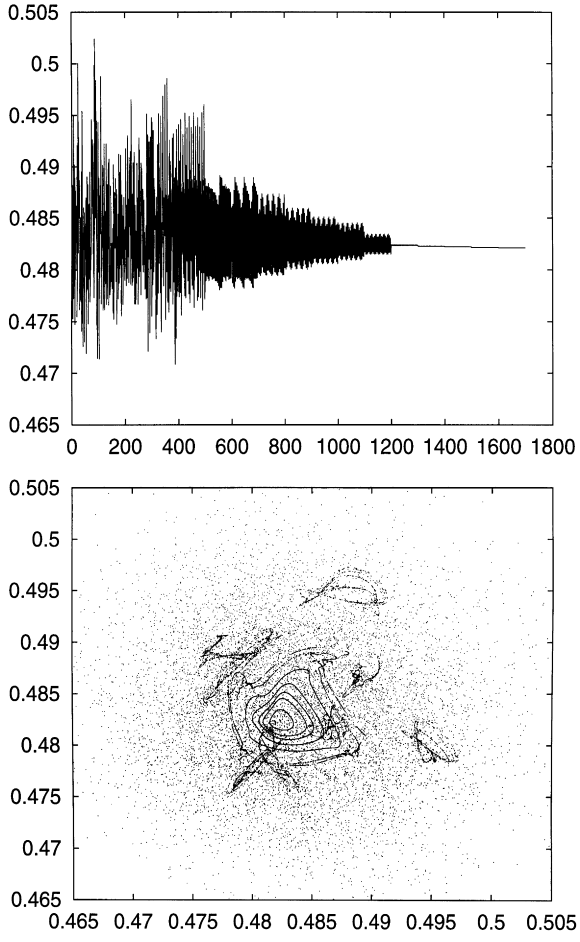

Fig. 6. (A) Evolution of the stimulus-forced dynamics $m_{\text{net}}(t)$ versus time t during the learning process. There is one learning step every 100 time step. In that example the dynamics is reduced to a fixed point after 12 learning steps ($t = 1200$). (B) Representation of the corresponding attractors in the $[m_{\text{net}}(t), m_{\text{net}}(t+1)]$ phase plane. Note the decrease of the diameter of limit cycles at each learning step.

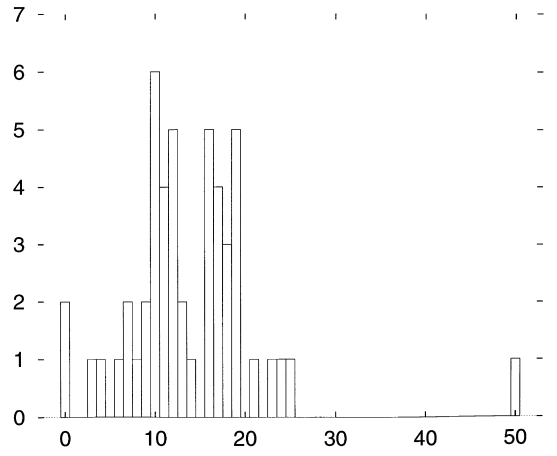

Fig. 7. Repartition of the number of the learning steps which have been necessary to reduce the dynamics to a fixed point. Simulation based on 50 learning processes completed on 50 different networks.

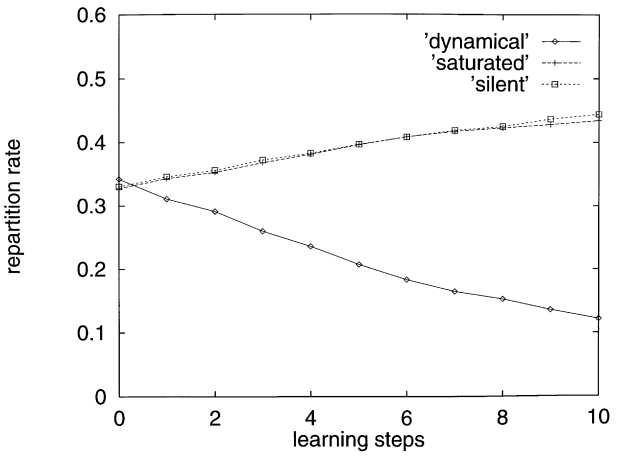

Fig. 8. Variation of the dynamical repartition of neurons during the 10 first steps of the learning process, averaged over 50 networks.

We generate random stimuli with $\sigma_I = 0.7$. For this value, the MFT predicts a chaotic behavior (see Fig. 2). However, for a finite size network, the occurrence of a stable stationary regime (2 torus, periodic attractor or fixed point) is observed in some cases and the probability of this occurrence cannot be neglected (see Table 2). We observe that with the previously mentioned parameters values, the stimulus-forced dynamics before learning is chaotic for 44 examples out of the 50 simulations we achieved (88% of the case). For a given finite-size network, with a random stimulus of given statistical characteristics, the probability for the stimulus-forced dynamics to be non-chaotic is called the reactivity of the network. The different reactivities of the 50 networks we tested are represented on Fig. 10. The average empirical reactivity over a sample of 50 networks is 12%.

Then we implement the learning rule. The value of the learning rate for the simulations is $\alpha = 0.1$. Smaller learning rates have been tested without any difference apart slowing down of the reduction of dynamics. With that learning rate, the modifications of the weights values are slight—about 0.6% per weight at each learning step.

The asymptotic dynamics is observed for each dynamical combination (network + stimulus) and for each learning step after discarding a transient regime of 500 iterations. In Fig. 6 we display the result of a typical simulation.

Fig. 6 shows the variations of the stimulus-forced dynamics of the network with the plot of $m_{\text{net}}(t)$. We begin learning at $t = 100$. Then we have one learning step every 100 step time. We stop the learning process when the dynamics are at a fixed point.

So, while the learning process is going on, we observe that the *attractor nature changes according to an inverse quasi-periodic route* from a strange attractor to a 2-torus, a limit-cycle and eventually a fixed point attractor. This phenomenon of reduction of dynamics has always been observed (Quoy et al., 1995).

Note that the change of the weights induces a reduction only for the stimulus-forced dynamics, which is specific of the association between *that* network and *that* stimulus. In that sense, the network has learned the stimulus. The network systematically reacts (reduces its dynamics) when that stimulus is presented, and the presentation of another random stimulus does not induce systematic reactivity (see Section 4.2 hereafter). Especially after the completion of learning, the autonomous dynamics (dynamics without stimulus) remains chaotic for all the cases.

However, learning efficiency is subject to variations due to the specific association network realization/stimulus realization. Fig. 7 gives the statistics of learning steps that have been necessary to reduce the dynamics of each of the 50 networks which have been tested. The mean number of learning steps is 14.

Note that for two associations network/stimulus, the number of learning steps is zero. This means that no learning was necessary because of a ''spontaneous'' reactivity between that network and that stimulus before learning. Those networks did not need to learn something they already ''knew''.

The second-order statistics of the connection weights are not strongly modified. According to our measures, the weights remain gaussian, and some slight changes are observed for mean and variance (Quoy, 1994). For instance, we observe that for each neuron, the sum of its afferent weight $W_i = \sum_j J_{ij}$ is slowly growing in absolute value at each learning step. The increase is about 0.006 at each learning step, for $\alpha = 0.1$, which means 0.6%.

Actually, this little change is linked to a significant change in the repartition of neural activation. The population of what we called ''dynamical neurons'' Section 3.2 is quickly moving towards extinction (see Fig. 8).

This can be explained in the following way. At each learning step, the weights modifications are slight enough to ensure that the dynamical repartition of neurons is not disrupted. In particular, most of the saturated neurons remain saturated while learning. As we shall see, this is a

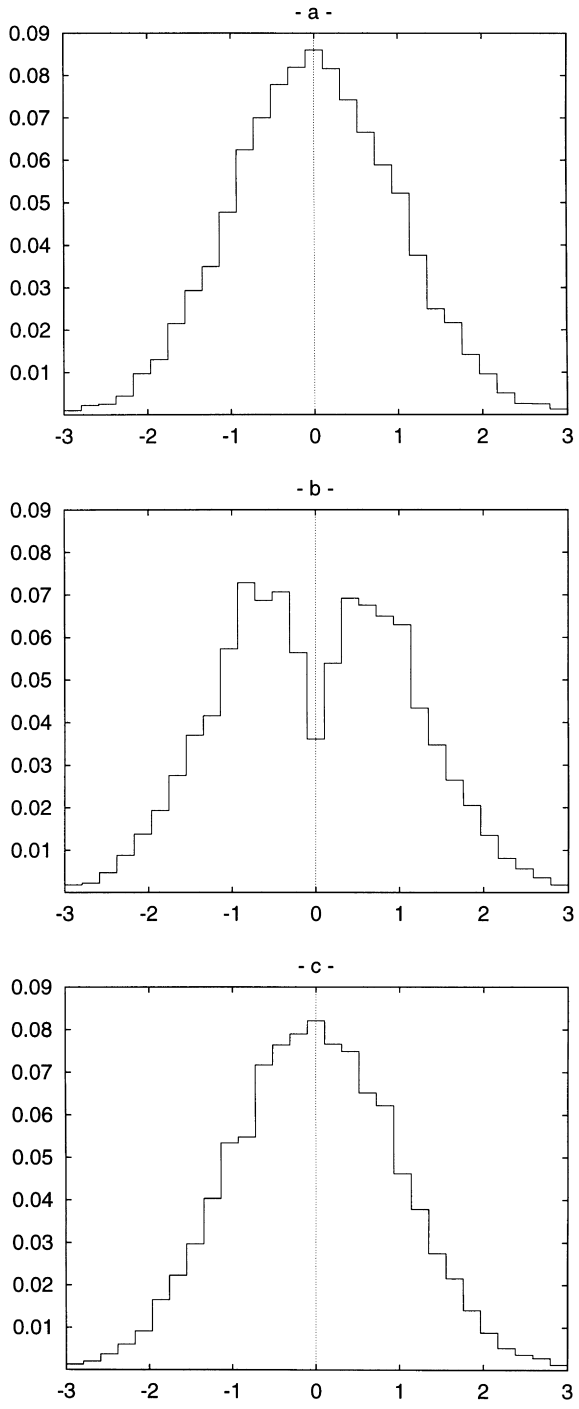

Fig. 9. Empirical distribution of the mean local field u_i^*. (A) Distribution of u_i^* before learning for the stimulus-forced dynamics. (B) Distribution of u_i^* after learning for the stimulus-forced dynamics. (C) Distribution of u_i^* after learning, with the dynamics induced by another random non-learned stimulus. The figures are the average histograms over 50 network histograms.

major key for the understanding of the learning process. *The system learns by reinforcing the influence of saturated neurons through synaptic weights*:

- if $x_i(t+1) > 0.5$, then $\Delta J_{ij} \geqslant 0, \forall j$. In this case the afferent synaptic weights increase. So the saturated neurons $[x_j(t) > 0.99]$ contribution increases which implies an increase of the target neuron's input. The target neuron is driven from the dynamical category to the saturated one.
- Reversely, if $x_i(t+1) < 0.5$, then $\Delta J_{ij} \leqslant 0, \forall j$. The saturated neurons $(x_j(t) > 0.99)$ contribution decreases and the target neuron is driven from the dynamical category to the silent one.

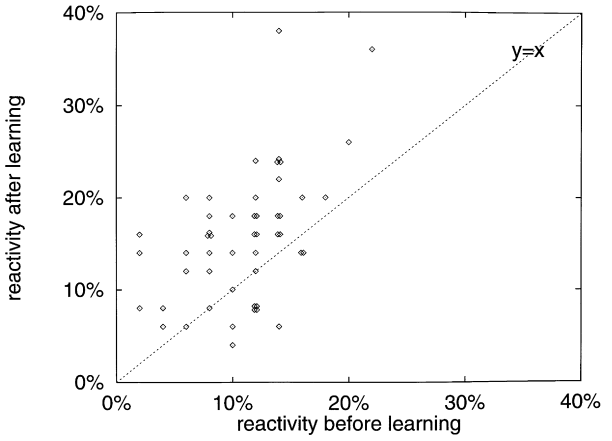

Fig. 10. Reactivity to random stimuli on a population of 50 networks before and after learning, with $g = 15$.

The ''effective threshold'' (see Section 3.2), which includes the input coming from saturated neurons, *grows in absolute value* for every neuron, so that one can simply say that learning reinforces the effective threshold *as if* σ_I *was increasing*.

So the changes in the dynamics are drastic. Actually, as soon as the learning process is performed, the MFT is not valid any more. Hereafter we will show the observed changes in the mean local field u_i^* distribution (see Section 2.2), which is closely linked to neural dynamical organization[1].

Fig. 9 gives us an illustrative overview of the learning process main actions:

- Strong modification of neural dynamical repartition.
- No more MFT predictability. The gaussian shape of the u_i^* distribution is lost.
- Specificity of the network's reactivity to the learned stimulus.

According to the MFT, the distribution of neural mean local fields is a normal law. This is verified quite accurately for finite-sized networks in the case of a stimulus-forced dynamics (Fig. 9A). Moreover, it is one basic assumption of the ''propagation of chaos'' property. Then learning the stimulus induces a strong modification of the distribution (Fig. 9B). Indeed, after learning, the mean local field u_i^* distribution is clearly non gaussian, so the characteristics

[1] A neuron with a strongly positive mean local field is saturated, and a neuron with strongly negative mean local field is silent.

of the ''propagation of chaos'' regime are quickly disappearing. Moreover, this qualitative modification is *specific* of the learned stimulus-forced dynamics. For any other non-learned stimulus-forced dynamics, the distribution of u_i^* remains gaussian (Fig. 9c).

4.2. Specificity of the stimulus sensitization

In fact, one can speak of learning if the network has a selective response for the learned stimulus, and if the learning procedure does not affect the dynamical behavior when another stimulus is presented.

In order to study this selectivity property, we have compared the reactivity of 50 networks to 50 random stimuli ($\bar{I} = 0, \sigma_I = 0.7$) before and after learning. In Fig. 10, each point represents the evolution of the network's reactivity.

So we see that global reactivity is increased slightly by learning: 12% before learning, 16% after learning, with a high variability from one network to another. Note that the two upper outlayers correspond to long learning processes (50 learning steps for upper left and 21 learning steps for upper right). Overlearning is dangerous for the specificity of the reactivity.

4.3. Presentation of a noisy stimulus

In order to study the robustness of the learning scheme, we add noise to a previously learned stimulus, and show how it affects the reactivity. We add to each component a gaussian noise of mean zero and standard deviation 0.07 or 0.14, which thus corresponds to a level of noise of 10% or 20%. 50 noisy stimuli have been presented to each network before and after the completion of the learning process. Results are in Fig. 11.

The mean reactivity of the 50 networks to noisy versions of the learned stimuli is 87% for a level of noise of 10% and 76% for a level of noise of 20%. Fig. 11 (with a 20% noise) illustrates the variability of the learning process efficiency from one network to another. Note that the points which are

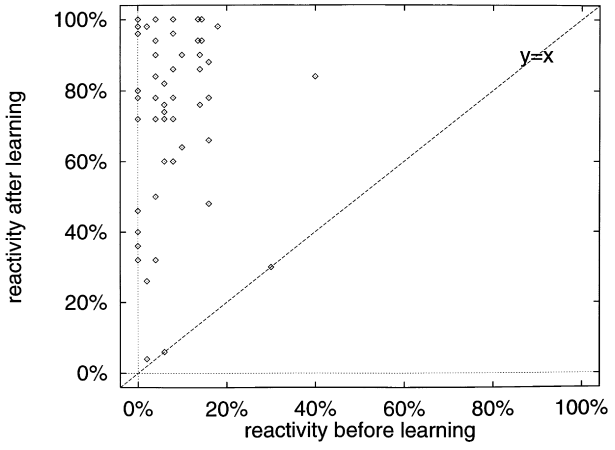

Fig. 11. Reactivity to noisy versions (noise = 20%) of the learned stimulus on a population of 50 networks before and after learning, with $g = 15$.

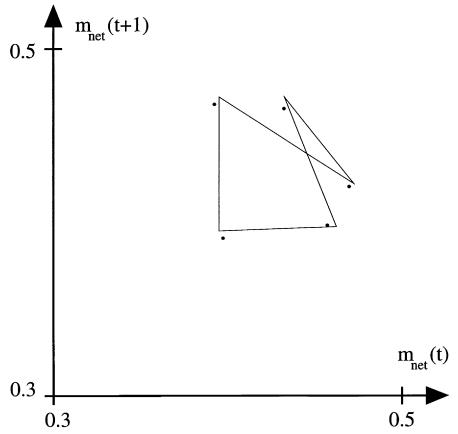

Fig. 12. The two attractors: 1st is associated to the learned stimulus (dots), 2nd to a noisy stimulus (dots linked by lines)[2].

on the ($y = x$) line correspond to networks for which we didn't iterate the learning process because of a spontaneous reactivity to the stimulus before learning (first histogram column on Fig. 7). Moreover, the points which are close to the ($y = x$) line correspond to short learning processes (about five learning steps). Conversely, long learning processes induce a very strong reactivity to noisy versions of the learned stimulus. So, in regards to Section 4.2, there is a necessary compromise between efficiency and specificity.

Reactivity can be viewed as the most simple cognitive operation, and means only a yes/no answer. We tried hereafter to perform a more elaborate answer—recognition of the learned stimulus based on limit cycles comparison. *The following measures are exploratory*.

In order to perform *recognition*, we now take as stimulus a vector of gaussian random variables with mean zero and standard deviation 0.1. We iterate the learning rule on one network until a *limit cycle* is reached. We add a gaussian noise of mean zero and standard deviation 0.01 (10%), and 0.02 (20%) to each component.

We recall that the presentation of a noisy stimulus changes the parameters of the system. Therefore, the attractor reached will be different from the one associated to the non noisy stimulus. Unlike other models where a noisy stimulus changes the initial conditions of the system, and where the recognition is performed when these initial conditions are in the attraction basin associated to the learned stimulus, here we have to deal with different systems for each noisy stimulus. So recognition has to be defined by some similarities between the attractors. These similarities can be observed very often in simulations. However, there are finite size effects and variations among individuals.

In order to quantify the similarity between cycles, we compute their gravity center, their mean radius and winding number. We take one of the mean radius as reference and

[2] In that figure, due to frequence locking, the dynamics of the networks goes to only five different values as soon as the permanent regime is reached. This is a small size effect.

Table 3
Ratio between the distance between the gravity centers and the reference radius (r_G), ratio between the two mean radius (r_R) and difference between the winding numbers (ρ) of the cycles of Fig. 12

rG	*rR*	ρ
0.0584	0.9358	0.027

then calculate the ratio between the distance between the gravity centers and the reference radius (r_R), the ratio between the two mean radius (r_G) and the difference between the winding numbers (ρ). Hence, we can heuristically define when two cycles are similar (i.e. coding for the same stimulus). The criterion is: $r_G < 0.15$ and $0.85 < r_R < 1.15$ and $\rho < 0.05$.

The values were established on a panel of randomly generated limit cycles with various frequencies, radius and centers. Table 3 gives the similarity values for the cycles displayed Fig. 12 (comparison of the attractor of the prototype learned stimulus, and the attractor of the prototype noised stimulus). We say that two stimuli are the same when the two attractors corresponding to them meet the previous criterion. Reciprocally, when this criterion is not fulfilled, the stimuli are said to be different. Hence recognition is based on this criterion.

To estimate the recognition rate, we present 30 noisy stimuli (derived from the prototype one) and compute the similarity values. Table 4 displays the recognition percentage after seven and 10 learning steps. Recognition is slightly increased by additional learning. Noise affects the recognition performance.

5. Discussion

The present paper uses MFT to show that the presentation of external stimulus to RRNN induces self-organization and dynamical reduction. This reduction comes from a nonlinear and collective effect. When a stimulus is presented, some neurons are driven out of the dynamical regime by the modification of their local field. Then it is shown that a learning procedure encodes the dynamical reduction in the weights. With slight changes that do not affect the global statistics of the weight values, the network is prepared to receive the stimulus and to react against its presentation. It has been shown that this increased reactivity is specific to the learned stimulus and its noisy versions.

Table 4
Recognition percentage versus the variance of the noise and the number of learning steps for a particular stimulus

	Level of noise	
	10%	20%
7 steps	0.83	0.27
10 steps	0.87	0.27

An important property of our model is to take into account the dynamical aspects of recurrent systems. Recognition is not a static process (reaching of a stable fixed point), but a dynamical one—convergence onto a complex attractor. The system does not need to wait until a fixed point is reached to perform recognition. Relying on our experiments, convergence on an attractor is fast. Thus chaos gives a high response flexibility, that systems with fixed point convergence do not have.

Hence this model is coherent with the observations by Freeman concerning the dimension reduction of the system attractor by recognition of a learned pattern, in a model of the olfactory bulb (Freeman et al., 1988; Yao and Freeman, 1990). In particular, a very simple network like ours, without any geometry, spikes and delays, exhibits such behavior. This gives an insight into the mechanism leading to such a phenomenon by the extraction of the few relevant parameters related to it. First investigations suggest that the basic principles of the present work, *propagation of chaos for RRNN and dynamical reduction induced by hebbian learning*, are still true in a more general framework.

Indeed, it is necessary to overcome the limitations of homogeneous RRNN and to investigate more complex RRNN models. The interaction of several neural models and hierarchical architecture with interconnected modulus of densely connected subnetworks are interesting features which we are introducing at the present stage of research. So it will be possible to compare the behavior of theoretical RRNN and real-life nervous systems (Linster et al., 1993).

Coding by dynamical attractors is also particularly suited for the learning of temporal signals. For the moment, we only focused on the learning of static stimuli, mainly because they fit within a theoretical frame. However, we did some preliminary simulations on presentation of temporals sequences (Quoy et al., 1993). The problem of the arrival of external oscillatory inputs remains open and is currently being investigated. This could lead to the connection of different chaotic networks in order to perform recognition tasks using the synchronization processes highlighted by Gray et al. (1989).

This research has also been inspired by more applied considerations. Are dynamical attractors networks able to overcome the limitations of relaxation attractor networks which are designed to converge to fixed points? The flexibility of chaotic attractors has been outlined by several authors (Skarda and Freeman, 1987; Hirsch, 1989; Palmadesso and Dayhoff, 1995) and strange attractors are sometimes considered as a ''reservoir'' of a wide variety of attractors in order to design powerful associative memories. Our research reinforces the long-term interest of these considerations and their short-term incompleteness.

First, the idea of permanently submitting the input to the system leads to an attractor which is particular to the association network-stimulus. So the system keeps the memory of the input and that complicates the recognition task. Then

the ability to decide is more complex than when the system just stops and some kind of frequency-locking mechanism is probably necessary to achieve the decision process. Eventually, the variability of the finite-size effect shows that more complex architecture has to be tested to improve the reliability of the system.

So it appears that the RRNN models are still not a component of dynamical and versatile neuromimetic information-processing systems, but they are a necessary step towards the design of stimulus recognition distributed dynamical systems.

Acknowledgements

This research has been partly supported by DRET (R&D Department from French M.O.D.) and by the Cogniscience research program of the C.N.R.S. through PRESCOT, the Toulouse network of searchers in Cognitive Sciences. B. Cessac was supported by a Cognisciences grant of the C.N.R.S. E. Daucé is supported by a MRE grant of French M.O.E.

References

Alkon D.L., Blackwell G.S., Rigler A.K., & Vogl T.P. (1990). Pattern-recognition by an artificial network derived from neural biological systems. *Biological Cybernetics, 62,* 363–376.

Amari S.I. (1972). Characteristics of random nets of analog neuron-like elements. *IEEE Transactions Systems in Man and Cybernetics, 2(5),* 643–657.

Babloyantz A., Nicolis C., & Salazar J.M. (1985). Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters, 111A,* 152–156.

Ben Arous G., & Guionnet A. (1995). Large deviations for Langevin spin glass dynamics. *Probability Theory of Related Fields, 102,* 455–509.

Changeux J.P., & Dehaene S. (1989). Neuronal models of cognitive function. *Cognition, 33,* 63–109.

Cessac B. (1995). Increase in complexity in random neural networks. *Journal de Physique I, 5,* 409–432.

Cessac B., Doyon B., Quoy M., & Samuclides M. (1994). Mean-field equations, bifurcation map and route to chaos in discrete time neural networks. *Physica D, 74,* 24–44.

Derrida B. (1988). Dynamics of automata, spin glasses and neural networks models. In *Non Linear Evolution and Chaotic Phenomena*. Plenum: New York.

Doyon B., Cessac B., Quoy M., & Samuclides M. (1993). Chaos in neural networks with random connectivity. *International Journal of Bifurcation and Chaos, 3 (2),* 279–291.

Eckhorn R., Bauer R., Jordan W., Brosch M., Kruse W., Munk M., & Reitboeck H.J. (1988). Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analysis in the cat. *Biological Cybernetics, 60,* 121–130.

Freeman W.J., Yao Y., & Burke B. (1988). Central pattern generating and recognizing in olfactory bulb: a correlation learning rule. *Neural Networks, 1,* 277–288.

Gallez D., & Babloyantz A. (1991). Predictability of human EEG: a dynamical approach. *Biological Cybernetics, 64,* 381–392.

Geman S. (1982). Almost sure stable oscillations in a large system of randomly coupled equations. *SIAM Journal of Applied Mathematics, 42,* 695–703.

Gray C.M., Koenig P., Engel A.K., & Singer W. (1989). Oscillatory responses in cat visual cortex exhibit intercolumnar synchronisation which reflects global stimulus properties. *Nature, 338,* 334–337.

Ginzburg I., & Sompolinsky H. (1994). Theory of correlations in stochastic neural networks. *Physics Review E, 50,* 3171–3191.

Hebb D. (1949) *The Organization of Behavior*. New York: Wiley.

Hirsch M.W. (1989). Convergent activation dynamics in continuous time networks. *Neural Networks, 2,* 331–349.

Hopfield J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Science USA, 79,* 2554–2558.

Linster C., Masson C., Kerszberg M., Personnaz L., & Dreyfus G. (1993). Computational diversity in a formal model of the insect olfactory macroglomerulus. *Neural Computation, 5,* 228–241.

Molgedey L., Schuart J., & Schuster H.G. (1992). Suppressing chaos in neural networks by noise. *Physics Review Letters, 69 (26),* 3717–3719.

Palmadesso P.J., & Dayhoff J.E. (1995). Attractor locking in a chaotic network: Stimulus patterns evoke limit cycles. *WCNN Washington, 1,* 254–257.

Quoy M., Cessac B., Doyon B., & Samuclides M. (1993). Dynamical behavior of neural networks with discrete time dynamics. *Neural Network World, 3(6),* 845–848.

Quoy M. (1994). *Apprentissage dans les réseaux neuromimétiques à dynamique chaotique*, PhD Thesis, ENSAE Toulouse.

Quoy M., Doyon, B., Samuelides, M., 1995. *Dimension reduction by learning in discrete time chaotic neural networks*, WCNN Washington 1.

Renals S., 1990. *Chaos in neural networks*, Eurasip Workshop, 90–99.

Skarda C.A., & Freeman W.J. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences, 10,* 161–195.

Sompolinsky H., Crisanti A., & Sommers H.J. (1988). Chaos in random neural networks. *Physics Review Letters, 61,* 259–262.

Sznitzman A.S. (1984). Non linear reflecting diffusion process and the propagation of chaos and fluctuations associated. *Journal of Functional Analysis, 56,* 311–336.

Yao Y., & Freeman W.J. (1990). Model of biological pattern recognition with spatially chaotic dynamics. *Neural Networks, 3,* 153–170.