
Natural Computing **3:** 135–157, 2004.

Short term memory in recurrent networks of spiking neurons

EMMANUEL DAUCÉ
UMR Movement and Perception, Faculty of Sport Sciences, University of the Mediterrannean, 163 avenue de Luminy, CP 910, 13288 Marseille cedex 9, France
(E-mail: dauce@esm2.imt-mrs.fr)

Abstract. We present in this paper a general model of recurrent networks of spiking neurons, composed of several populations, and whose interaction pattern is set with a random draw. We use for simplicity discrete time neuron updating, and the emitted spikes are transmitted through randomly delayed lines. In excitatory-inhibitory networks, we show that inhomogeneous delays may favour synchronization provided that the inhibitory delays distribution is significantly stronger than the excitatory one. In that case, slow waves of synchronous activity appear (this synchronous activity is stronger in inhibitory population). This synchrony allows for a fast adaptivity of the network to various input stimuli. In networks observing the constraint of short range excitation and long range inhibition, we show that under some parameter settings, this model displays properties of –1– dynamic retention –2– input normalization –3– target tracking. Those properties are of interest for modelling biological topologically organized structures, and for robotic applications taking place in noisy environments where targets vary in size, speed and duration.

Key words: chaotic dynamics, delays, neural field, random recurrent neural networks, short term memory, sparsely connected networks, spiking neurons, synchronization

1. Introduction

Neural networks are generally considered as information-processing systems, i.e. as systems which operate transformations of their inputs in order to produce outputs. This point of view tends to favour the study of feed-forward architectures which structurally obey to the information-processing paradigm. Under this assumption, a given input may always lead to the same output, with a fluctuation tolerance rate, which corresponds to the noisiness of the channel (Shannon and Weaver, 1949). This approach is of course efficient when dealing with classification systems, identification systems, and some control systems, under which the parameter setting of the network (learning) leads to the best signal/noise rate in processing the output. This approach is however bound to encounter serious difficulties as soon as one wants to deal with *real* temporal processing, for instance classifying some

perceptual events that take place under various places, speed and duration, as in a video flow, switching attention, coordinating several perceptual fields, following intermittent targets, etc. ... Even if some feed-forward solutions can be designed for every particular problem, there is no unifying framework to the question of processing signals that take place in space and time.

On the contrary, dynamical systems theory offers a natural framework for tackling the question of time and space variability in interaction processes. We can briefly summarize them as:

- The state space of a dynamical system is divided into several *basins of attraction* into which every trajectory converges towards a unique *attractor*.
- The nature and spatio-temporal pattern of these attractors can vary strongly from one system to the other, and from one basin of attraction to the other. The most simple attractors are fixed points. The most complex attractors are called "strange attractors". They are associated with chaotic dynamics.
- The shape and nature of attraction basins can vary under parametric changes, which are often supposed adiabatic (i.e. slow according to the state update). Phase transitions (or bifurcations) occur when the topological characteristics of an attractor basin undergoes a sudden change.

In consequence, a dynamical system can display two sorts of temporal phenomena. First, the trajectory displays the short term causality between the successive states of the system. Second, the evolving topology of its attraction basins display some structural changes occurring at discrete moments. The history of those bifurcations constitutes the *slow dynamics* (i.e. the history of "catastrophes").

The brain is known to be a strongly recurrent structure (brain structures at different scales are in majority composed of re-entrant connections, and the connections between different structures/areas are always reciprocal, see for instance (Hupé et al., 1998) for the visual cortex). This massive recurrence suggests a major role of self-feeding dynamics in the processes of perceiving, acting, and learning, and more generally in maintaining the organism alive. We thus want to identify some of the advantages of re-entrant activity in natural perception and action processes, and thus use a recurrent neural system as a model of natural perception.

2. Dynamic models

In the field of neural modelling, spike coding has increased in popularity at the expense of rate coding, even if they in some respects may correspond to

the same coding through different points of view (Gerstner, 1999). At first regard, the qualification of a neuron model as spiking or rate coding model seems to rely on the nature of the neuron output as discrete or continuous, so that McCullogh and Pitts model (Mac Cullogh and Pitts, 1943) may for instance be considered as a spiking model. However, the core of the distinction between the two families of networks relies on the way analog signals may be transmitted, either through a continuous non-linear output, or through a *temporal* order or phase shift. The spiking approach basically means introducing *time* into the model.

Introducing time, or "dynamics", into the model, has several meanings. The first is to introduce dynamics into the local model of neuron, so that its output at a given time is dependent on past local events, such as recent firing or recent PSP (Post Synaptic Potential) arrival. This is typically the case in the Hodgkin-Huxley model, the "integrate and fire" models (Gerstner, 1999) as well as in explicit dynamic neuron models (Izhikevich, 2003). The other meaning of dynamic is to take into account the delays of transmission between cells, so that one can for instance explicitly encode into the synaptic weights a chain of synchronous firing patterns (Tsodyks and Feigelman, 1988; Herz et al., 1989; Hermann et al., 1995). The formal correspondence with (discrete-time or continuous) dynamical systems is obtained when the model owns one or several loops in the connectivity pattern, when the transmission time is explicitly modelled, either through axonal or cellular mechanisms, and when the spike emission is noiseless, see for instance (Hansel and Sompolinsky, 1996).

2.1. *A generic recurrent system*

We take for simplicity a model with discrete time updating, so that spike emissions are bounded to take place on a discrete temporal scale (the time unit is of the order of one millisecond). The transmission times are explicitly encoded between every neuron with a discrete positive value (those transmission times may correspond to the sum of an axonal delay and PSP transmission delay through the dendrites). The connectivity pattern is sparse.

Our model is a discrete-time dynamical system with parallel update, where the state of the system at time t both depends on the state of the system at time $t-1, t-2, t-3$ (etc. ...) and on the input $\mathbf{u}(t)$, i.e.

$$\mathbf{x}(t) = f(\mathbf{x}(t-1), \ldots, \mathbf{x}(t-\tau_{\max}), \mathbf{u}(t)) \tag{1}$$

where f is the global operator described in the following equations and $\tau_{\max}$ (maximal transmission delay) determines the degree. One can notice that our system is deterministic as soon as the input signal is set according to a deterministic process.

A network is defined as a pool of P interacting populations of neurons, of respective sizes $N^{(1)},\dots,N^{(P)}$. The global number of neurons is $N = \sum_{p=1}^{P} N^{(p)}$. The synaptic weights from population q towards population p are stored in a matrix $\mathbf{J}^{(pq)}$ of size $N^{(p)} \times N^{(q)}$ (possibly sparse). The state vector of population p at time t is $\mathbf{x}^{(p)}(t)$, of size $N^{(p)}$. The initial conditions $x_i^{(p)}(0)$ are set according to a random draw in $\{0, 1\}$.

At each time step $t \geq 1$, $\forall (p, q) \in \{1,\dots,P\}^2$, $\forall i \in 1,\dots,N^{(p)}$,

$$h_i^{(pq)}(t) = \sum_{j=1}^{N^{(q)}} J_{ij}^{(pq)} x_j^{(q)}(t - \tau_{ij}^{(pq)}) \tag{2}$$

is the *local field* of population q towards neuron i of population p, and $\tau_{ij}^{(pq)}$ is a transmission delay between neuron j and neuron i.

We also consider spatio-temporal input signals $\mathbf{u}^{(p)} = \{\mathbf{u}^{(p)}(t)\}_{t=1\dots+\infty}$, where $\mathbf{u}^{(p)}(t)$ is a $N^{(p)}$ dimensional input vector at time t on population p. The input $\mathbf{u}^{(p)}(t)$ acts like a bias on each neuron.[1] Then, the global equation of the dynamics is:

$$\begin{aligned} &\forall t \geq 1, \forall p \in \{1,\dots,P\}, \forall i \in \{1,\dots,N^{(p)}\} \\ &x_i^{(p)}(t) = H\left(-\theta_i^{(p)} + u_i^{(p)}(t) + \sum_{q=1}^{P} h_i^{(pq)}(t)\right) \end{aligned} \tag{3}$$

where H is a Heaviside function so that the neuron output takes its values in $\{0, 1\}$. Its activation potential has a real value, which corresponds to a linear combination of afferent local fields and input minus activation threshold $\theta_i^{(p)}$.

2.1.1. *The transmission delays*

The use of transmission delays in neural processing is traditionally associated to the problem of temporal sequences learning within recurrent dynamical networks. It has been shown for instance that Hebbian learning of temporal patterns remains efficient within a system containing a broad range of delays (Herz et al., 1989). The delay scattering is not a drawback for learning sequences with long term dependencies. On the contrary, a simple Hebbian mechanism allows to select the appropriate delayed line for the learning of a specific temporal dependency. Another important property of recurrent systems with delayed lines is the formation of "itinerant", sometimes chaotic dynamics, within systems storing spatial or spatio-temporal patterns as limit-cycle attractors (Riedel et al., 1988; Nützel et al., 1994; Hoshino et al., 1982).[2]

The use of delays in our model is a way to characterize the temporal relationships between populations of neurons, in particular for allowing different reaction times between populations of excitatory and inhibitory neurons. The pattern of delay transmission is non-homogeneous: it is set according to a a random draw, so that every delay $\tau_{ij}^{(pq)}$ is set according to

$$\tau_{ij}^{(pq)} = \tau_0^{(pq)} + p_{ij}^{(pq)}$$

where $\tau_0^{(pq)}$ is the minimal delay and $p_{ij}^{(pq)}$ is set according to a Poisson draw of parameter $\lambda^{(pq)}$. So, for every pattern of connection between two populations, there is typical transmission delay given by $\tau^{(pq)} = \tau_0^{(pq)} + \lambda^{(pq)}$, with a certain variability from one axon to the other. In our simulations, typical mean delays are between 3 and 8 time steps, so that maximal delays are of the order of 20–25 time steps (and minimum is 1 time step). As we will see in the following section, the use of non homogeneous delay patterns strongly favours the appearance of both complex and synchronized dynamics.

2.1.2. *Weights setting*

Our network belongs to the category of Random Recurrent Neural Networks (RRNNs), so that the weights obey to a random draw. The principal consequence of that setting is the "almost sure" non-symmetry of the connectivity pattern, so that one can not ensure the convergence of the dynamics towards a fixed point attractor. Autonomous RRNN's (i.e. $\forall t, \mathbf{u}(t) = 0$ in eq.(1)) are discrete time dynamical systems, that can for instance display a generic quasi-periodicity route to chaos while progressively increasing the gain of a continuous transfer function (Doyon et al., 1993).

Each family of weights $(J_{ij}^{(pq)})$'s (weights from population q towards population p) has a certain sparsity $\rho^{(pq)}$, so that a given weight $J_{ij}^{(pq)}$ is equal to 0 with probability $1 - \rho^{(pq)}$, and takes its value randomly in interval $\left[\frac{\bar{J}^{(pq)}}{\rho^{(pq)}N^{(q)}} - \sqrt{\frac{3}{\rho^{(pq)}N^{(q)}}}\sigma_J^{(pq)}, \frac{\bar{J}^{(pq)}}{\rho^{(pq)}N^{(q)}} + \sqrt{\frac{3}{\rho^{(pq)}N^{(q)}}}\sigma_J^{(pq)}\right]$, with probability $\rho^{(pq)}$, i.e.

$$J_{ij}^{(pq)} = \begin{cases} \frac{\bar{J}^{(pq)}}{N_{\text{aff}}^{(pq)}} + \frac{\sigma_J^{(pq)}}{\sqrt{N_{\text{aff}}^{(pq)}}} b & \text{with probability } \rho^{(pq)} \\ 0 & \text{with probability } (1 - \rho^{(pq)}) \end{cases} \tag{4}$$

where

- $N_{\text{aff}}^{(pq)} = \rho^{(pq)}N^{(q)}$ represents the expectation of the number of afferent weights arriving from population q.
- b is set according to $\mathcal{U}(0, 1)$, which is a uniform distribution in $[-\sqrt{3}, \sqrt{3}]$

Those settings are such that the expectation of the weights is $E\left(J_{ij}^{(pq)}\right) \simeq \frac{\bar{J}^{(pq)}}{N^{(q)}}$ and $\text{var}\left(J_{ij}^{(pq)}\right) \simeq \frac{\left(\sigma_{\text{J}}^{(pq)}\right)^2}{N^{(q)}}$.

In this paper, in order to remain coherent with elementary biological requirements, we rule the weights of a given population $q \in \{1,\dots,P\}$ to be either purely excitatory or purely inhibitory. This constraint implies additional dependencies between $\bar{J}^{(pq)}$, $\sigma_{\text{J}}^{(pq)}$, $N^{(q)}$ and $\rho^{(pq)}$. For a given couple $(\bar{J}^{(pq)}, \sigma_{\text{J}}^{(pq)})$, we note $d^{(pq)} = \frac{\bar{J}^{(pq)}}{\sigma_{\text{J}}^{(pq)}}$, which represents the eccentricity of a weight distribution. One can determine the maximal number of afferent links $N_{\max}^{(pq)}$ such that $\frac{\bar{J}^{(pq)}}{N_{\max}^{(pq)}} - \sqrt{\frac{3}{N_{\max}^{(pq)}}}\sigma_{\text{J}}^{(pq)} = 0$, i.e. $N_{\max}^{(pq)} = \frac{\left(\bar{J}^{(pq)}\right)^2}{3\left(\sigma_{\text{J}}^{(pq)}\right)^2} = \frac{(d^{(pq)})^2}{3}$. When $N^{(q)}$ is set, the maximal sparsity is thus $\rho_{\max}^{(pq)} = \frac{N_{\max}^{(pq)}}{N^{(q)}} = \frac{(d^{(pq)})^2}{3N^{(pq)}}$.

We also define parameter $a^{(pq)}$ such that every $J_{ij}^{(pq)}$ is bound to the interval $[0, \frac{a^{(pq)}}{N_{\max}^{(pq)}}]$. This parameter represents the expanse of afferent weights interval.

For given $a^{(pq)}$ and $d^{(pq)}$, the values of $\bar{J}^{(pq)}$ and $\sigma_{\text{J}}^{(pq)}$ are fixed i.e.:

$$\bar{J}^{(pq)} = \frac{a^{(pq)}}{2}, \quad \sigma_{\text{J}}^{(pq)} = \frac{a^{(pq)}}{2d^{(pq)}}$$

Or, reversely, setting $\bar{J}^{(pq)}$ and $\sigma_{\text{J}}^{(pq)}$ fixes $a^{(pq)}$ and $d^{(pq)}$:

$$a^{(pq)} = 2\bar{J}^{(pq)}, \quad d^{(pq)} = \frac{\bar{J}^{(pq)}}{\sigma_{\text{J}}^{(pq)}}$$

2.1.3. *Activation thresholds*
The activation threshold $\theta_i^{(p)}$'s are randomly set according to the Gaussian distribution $\mathcal{P}(\bar{\theta}^{(p)}, \sigma_\theta^{(p)})$.

2.1.4. *Strong sparsity*
In case sparsity $\rho^{(pr)} = \rho_{\max}^{(pr)} = \frac{\left(\bar{J}^{(pr)}\right)^2}{3\left(\sigma^{(pr)}\right)^2 N^{(r)}} = O\left(\frac{1}{N^{(r)}}\right)$, the mean number of afferent links is $N_{\max}^{(pr)} = \frac{\left(\bar{J}^{(pr)}\right)^2}{3\left(\sigma^{(pr)}\right)^2}$.[3] The mean number of afferent links is thus independent of $N^{(r)}$.

For large $N^{(r)}$, the variance of the distribution of afferent connexions can not be neglected, from one neuron to the other. If we remind that $E(J_{ij}^{(pr)}) = \frac{\bar{J}^{(pr)}}{N^{(r)}}$ and $\mathrm{var}(J_{ij}^{(pr)}) = \frac{\sigma_J^{(pr)2}}{N^{(r)}}$, we must set $\bar{J}_0$ and σ_0 knowing that

$$\rho_0 = \frac{\left(\bar{J}^{(pr)}\right)^2}{3\left(\sigma_J^{(pr)}\right)^2 N^{(r)}}$$
$$J_{ij}^{(pr)} = \begin{cases} \frac{\bar{J}_0}{\rho_0 N^{(r)}} + \frac{\sigma_0}{\sqrt{\rho_0 N^{(r)}}} b & \text{with probability } \rho_0 \\ 0 & \text{with probability } (1-\rho_0) \end{cases} \tag{5}$$

where b is set according to $\mathcal{U}(0,1)$, which is a uniform distribution in $[-\sqrt{3}, \sqrt{3}]$.

So that

$$E(J_{ij}^{(pr)}) = \rho_0 \frac{\bar{J}_0}{\rho_0 N^{(r)}} = \frac{\bar{J}^{(pr)}}{N^{(r)}} \Leftrightarrow \bar{J}_0 = \bar{J}^{(pr)}$$

and

$$\begin{aligned} \mathrm{var}(J_{ij}) &= \rho_0(1-\rho_0)\left(\frac{\bar{J}_0}{\rho_0 N^{(r)}}\right)^2 + \rho_0\left(\frac{\sigma_0}{\sqrt{\rho_0 N^{(r)}}}\right)^2 \\ &= (1-\rho_0)\frac{3(\sigma_0)^2}{N^{(r)}} + \frac{(\sigma_0)^2}{N^{(r)}} = \frac{\left(\sigma_0\sqrt{4-3\rho_0}\right)^2}{N^{(r)}} \\ &\Leftrightarrow \sigma_0 \simeq \frac{\sigma_J^{(pr)}}{2} \text{ for small values of } \rho_0 \end{aligned}$$

If we set $\bar{J}_0 = \bar{J}^{(pr)}$, $\rho_0 = \frac{\left(\bar{J}^{(pr)}\right)^2}{3\left(\sigma_J^{(pr)}\right)^2 N^{(r)}}$, then $\sigma_0 = \frac{\sigma_J^{(pr)}}{\sqrt{4-3\rho_0}}$, then one can determine $\rho_1 = \frac{\left(\bar{J}^{(pr)}\right)^2}{3(\sigma_0)^2 N^{(r)}}$, and $\sigma_1 = \frac{\sigma_J^{(pr)}}{\sqrt{4-3\rho_1}}$, etc. ... This process can be iterated until it converges towards ρ^*, σ^*.

More simply, one can just remark that $\rho^* = (4-3\rho^*)\rho_0$, i.e. $\rho^* = \frac{4\rho_0}{1+3\rho_0}$, so that $\sigma^* = \frac{\sigma_J^{(pr)}}{\sqrt{4-3\rho^*}}$, and implement the weights according to $\bar{J}^{(pq)}$ and σ^* in order to reliably predict the behaviour of finite-size sparse networks.

2.2. *Excitatory-inhibitory interactions*

Most of the cortical an sub-cortical layers are composed of interacting populations of excitatory and inhibitory neurons. We present in this section a large scale structure. In order to take into account some basic physiological features, we suppose that our two-populations structure grossly models a cortical column, such that:

- local dynamics dominate incoming signals.
- excitatory neurons can receive and send signals, inhibitory neurons only act locally.
- excitatory neurons represent 70–90% of the total population.

The parameters setting globally defines the way the two populations interact. For the sake of simplicity, we suppose that each population of neurons is statistically homogeneous. The excitatory population is population 1. The inhibitory population is population 2. Inputs are only displayed on the excitatory layer, i.e. $\forall i, \forall t, u_i^{(2)}(t) = 0$ (see eq.(3)). We take binary inputs such that $\forall i, \forall t, u_i^{(1)}(t) \in \{0, 1\}$. The population sizes are set to $N^{(1)} = 1000$, $N^{(2)} = 200$.

We define two global parameters k and d which help to define the interaction pattern. The way the two populations interact are characterized by:

- The *asymmetry* "k" between excitatory and inhibitory influences. The higher is k, the more inhibitory weights dominate excitatory ones. The value $k = 1$ corresponds to a global balance between excitatory and inhibitory influences. Low values of k correspond to a dominant excitatory influence. Parameter k acts on the mean value of weights distributions such that:

$$\bar{J} = \begin{pmatrix} \frac{1}{2} & -\frac{k}{2} \\ \frac{k}{2} & -\frac{k}{2} \end{pmatrix}$$

- The *eccentricity* "d" of the weights distribution. This parameter mostly acts on the scattering/standard deviation of the weights such that:

$$\sigma_{\mathrm{J}} = \begin{pmatrix} \frac{1}{2d} & \frac{\sqrt{k}}{2d} \\ \frac{\sqrt{k}}{2d} & \frac{\sqrt{k}}{2d} \end{pmatrix}$$

The other network parameters relate to the threshold, i.e. to the individual excitabilities. We suppose that neurons are not far from the threshold, i.e. a small excitation can initiate the neuron to spike. The mean threshold is of the order of 1/10 of the expanse of the weights interval, i.e.

$$\bar{\theta} = \begin{pmatrix} 0.1 \\ 0.1k \end{pmatrix}$$

For the sake of simplicity, the threshold standard deviations are set to zero

$$\sigma_\theta = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Predictions on the *mean field* of such systems can be obtained in the limit of large sizes under an hypothesis of independence of the individual signals (Sompolinsky et al., 1988; Cessac, 1995), and under a condition of homogeneity of the distribution of the weights in a given population p. The different network regimes and their conditions of appearance are reliably predicted by the mean field equations. The validity of this approximation for large networks is out of the scope of this paper, see (Cessac, 1995; Moynot and Samuelides, 2002).

The mean field approach allows us to define some parametric regions under which different regimes can be obtained. One can for instance define a 2-dimensional parametric space, for which the first axis is the eccentricity d, and the second axis is the asymmetry k. In this space, one can define some frontiers between regions, some of them corresponding to quiescent activity, others to disordered activity, and others to synchronized activity. The exploration of the parametric space is not fulfilled at the present time. We have however identified some regions of interest under which the network may display interesting properties, like:

- A low level of activity (5 to 10% of the neurons may be active at a given time)
- Spontaneous chaotic dynamics (the chaotic activity allows for fast adaptivity towards incoming events)
- a strong reactivity towards input presentation

We present on Figure 1 the reaction of a network towards input presentation. In this simulation, the delays between neurons are strictly equal to 1. Once the system is defined, the dynamical system (2-3) is initialized with a binary random vector taking values in $\{0, 1\}$.[4] The spontaneous activity (without input) is displayed from $t = 1$ to $t = 100$. This activity is weak, irregular, with a slight synchronization (synchronous bursts appear as small peaks on the mean activity). The synchronization is stronger on inhibitory population. Then, an input is sent on 15 excitatory neurons. This input appears as a black band on first population activity. The network then tends to become more synchronous with almost periodic bursts of activity. This synchrony appears to be stronger in the inhibitory population. When the input is removed, the network turns back to its initial spontaneous weak activity.

The network dynamics appears to be very sensitive to its input, since a small input leads to a strong change in the nature of the displayed dynamics. The dynamical transition towards input presentation can be compared to a phase transition, i.e. the crossing of a bifurcation point in the parametric space. This kind of phase transition in input sensitivity has been observed for long time in neurobiology, see for instance (Skarda and Freeman, 1987), and also (Mac Leod and Laurent, 1996). More generally, synchronizing beha-

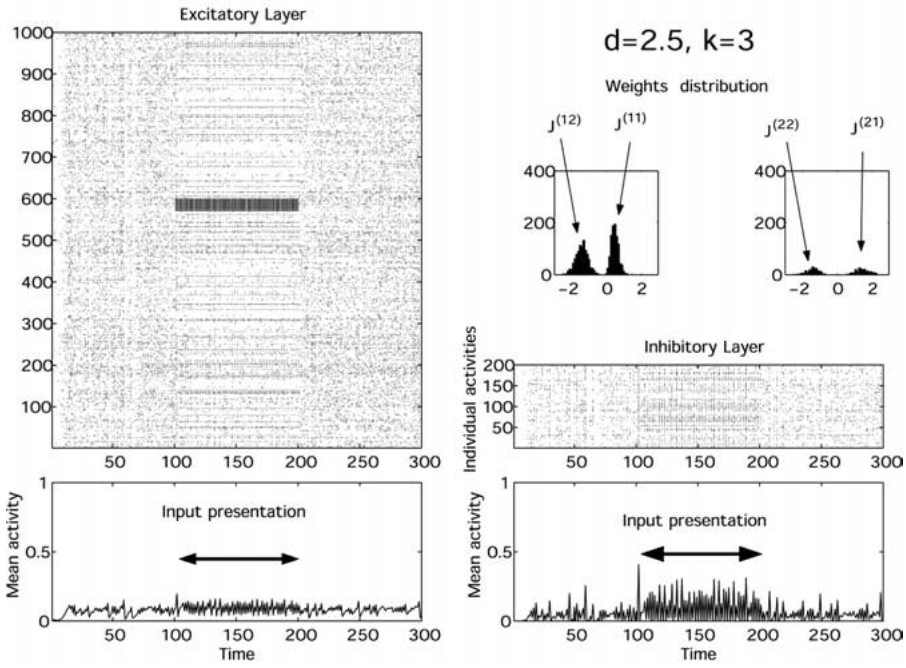

Figure 1. Input presentation. The network is composed of $N^{(1)} = 1000$ neurons and $N^{(2)} = 200$ neurons. The effective sparsity are $\rho^{(11)} = 0.8\%$, $\rho^{(12)} = 11\%$, $\rho^{(21)} = 2\%$, $\rho^{(22)} = 11\%$. An input is sent at time 101–200 on neurons 585–600. **Upper left**: global activity of the excitatory layer. **Lower left**: mean activity on excitatory layer. **Middle right**: global activity of the inhibitory layer. **Lower right**: mean activity on inhibitory layer. **Upper right insets**: distribution of the sum of incoming weights from excitatory and inhibitory populations, for 1000 excitatory neurons (left) and 200 inhibitory neurons (right).

viours in unitary delays networks depend on the asymmetry k, i.e. inhibition has to dominate excitation for the network to produce synchrony. This point has been overlooked in other simulation works, see for instance (Bush and Sejnowski, 1996; Izhikevich, 2003).

2.3. *Introducing delays*

A further step towards more realistic population simulation is to define a series of delays between the different populations. Effective delays depend on two factors: the axon conduction velocity and the neuronal integration time. Extra-cortical, intra-cortical and intra-columnar delays seem to be of the same order, despite the variety of axon length. Realistic excitatory delays seem to be of the order of 2–10 ms (Nowak and Bullier, 1997). The values of inhibitory delays are not so thoroughly documented. The role delays in

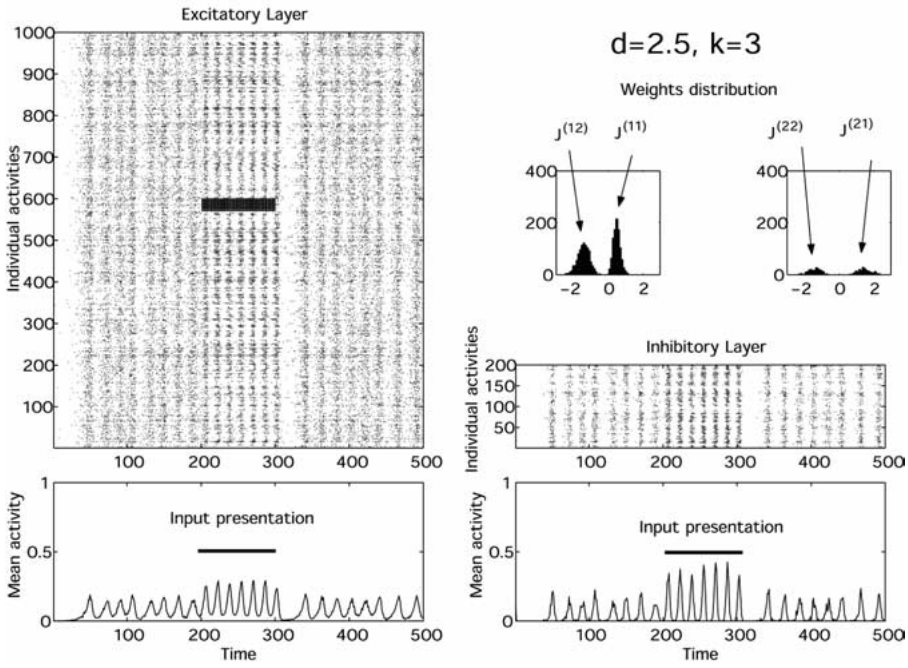

Figure 2. Input presentation. The network is composed of $N^{(1)} = 1000$ neurons and $N^{(2)} = 200$ neurons. Excitatory links are globally faster ($\simeq$ 3 time steps) than inhibitory links ($\simeq$ 5 time steps). An input is sent at time 201–300 on neurons 585–600. **Upper left**: global activity of the excitatory layer. **Lower left**: mean activity on excitatory layer. **Middle right**: global activity of the inhibitory layer. **Lower right**: mean activity on inhibitory layer. **Upper right insets**: distribution of the sum of incoming weights from excitatory and inhibitory populations, for 1000 excitatory neurons (left) and 200 inhibitory neurons (right).

neuron dynamics remain unclear, since simulation work suggests that high delays ($>$ 5 ms) may abolish synchronization (Bush and Sejnowski, 1996).

For our next simulation, the network parameters remain the same except we set the delays to random values. As mentioned (Bush and Sejnowski, 1996), we have noticed that synchronization and input reactivity tend to disappear when inhibitory and excitatory delays obey to the same distribution. We have thus set the delay parameters such that excitatory links are globally faster than inhibitory links:

$$\tau_0 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad \lambda = \begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix}$$

The minimal delay is thus 1. The mean delay is 3 for excitatory links, and 5 for inhibitory links. The resulting simulation is displayed on Figure 2.

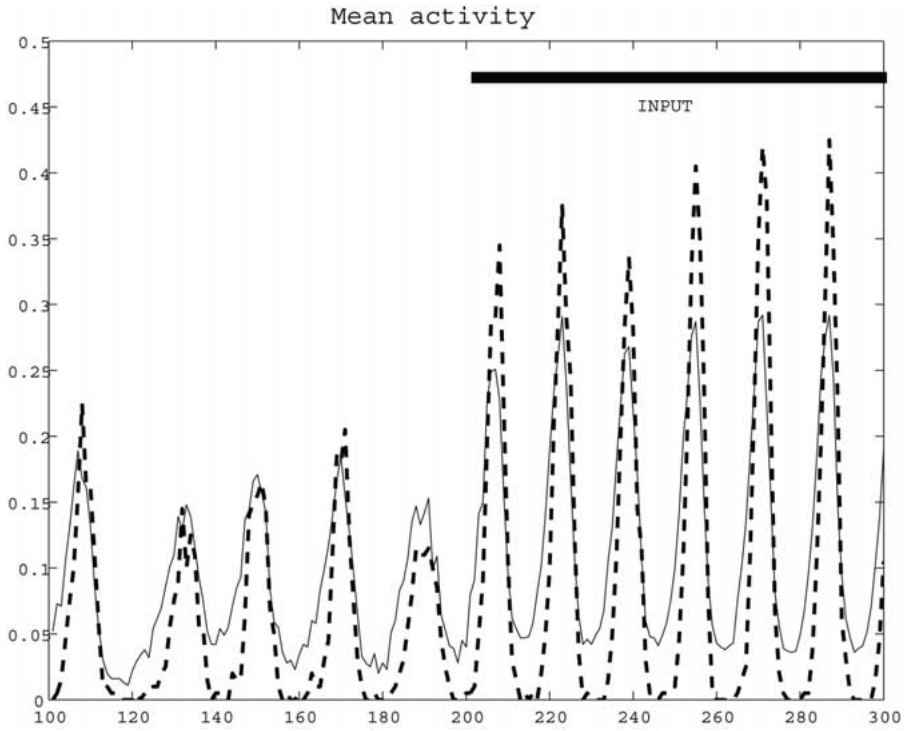

Figure 3. Mean activity between $t = 101$ and $t = 300$. The network is composed of $N^{(1)} = 1000$ neurons and $N^{(2)} = 200$ neurons. **Solid line**: mean activity of excitatory population. **Dashed line**: mean activity of inhibitory population.

The spontaneous activity of the network is now found to be synchronized, with irregular bursts of activity whose period is of the order of 20 time steps. The bursting behaviour does not rely on the intrinsic properties of neurons, like in (Izhikevich, 2003). The input presentation induces a transition towards stronger and more regular synchronization with a shorter period of the order of 15 time steps. When the input is removed, the network turns back to its initial spontaneous synchronized activity.

Figure 3 gives a zoom on the mean activities of excitatory and inhibitory populations between $t = 101$ and $t = 300$ (while an input is sent at $t = 201$). One can notice that excitatory population has a spontaneous tendency to increase its activity, which results in a higher activity of the inhibitory population, and the increase of inhibitory activity provokes a sudden decrease of excitatory activity few time steps later (and the cycle goes on). Sending an input lightly increases the mean level of excitatory population activity, which causes an increase of burst regularity and frequency. In both cases, synchronization relies on the late reaction of inhibitory population

(the inhibitory effect takes place 4–6 time steps after excitatory activity has significantly increased), which allows the excitatory population to maintain a strong spontaneous activity, even when no input is sent. This kind of slow synchronization is very robust, and reproducible for various networks and delays settings, provided *inhibitory delays are longer than excitatory ones* (in a 2/1 rate).

The very common observation of local synchrony in brain activity, especially in the gamma band (30–50 Hz), corresponds to a period (20–30 ms) which is 1 order higher than elementary delays (2–10 ms). This suggests that local synchronization processes could rely on a differential between excitatory and inhibitory delays.

3. Modelling short-term memory

The property of dynamic retention and its formal equivalence with short term memory in neural assemblies is another prototypical example of the natural connections between dynamical systems theory and neuroscience. In what follows, we will consider a particular family of topologically structured dynamical systems, and look at the basic requirements for such systems to manifest the property of dynamic retention. Topologically structured dynamical systems are defined by a *space* (or a *map*, which defines a distance), a *field*, possibly representing the states of uniformly distributed units, and a *process* by which those units interact. The nature of this interaction process is supposed to depend on the distance. A prototypic model of such topologically structured dynamical system is given by the reaction-diffusion equations of Alan Turing (Turing, 1952), which is a model of pattern formation in chemical reactions.[5]

Topologically structured dynamical systems as models of short term memory have been introduced in neural modelling with the Neural Field of Amari (Amari, 1977). This model approximates the neuron indexes (corresponding to a position in the map) as a continuous dimension. The original equation is the following:

$$\tau \frac{dh_i(t)}{dt} = -h_i(t) + u_i(t) - \theta + \int_{\ell \in \mathbb{R}} w(i-\ell) f\left(h_{i-\ell}(t)\right) d\ell \tag{6}$$

where i is a (continuous) index of position, h_i is a potential, u_i is an input, θ is an activation threshold, w is an interaction kernel, and f is the activation function. The parameter setting is mainly focused on the interaction kernel, which is a function of distance ℓ. The choice of a Mexican hat (second Gaussian derivative) shape function corresponds to a *short range excitation and*

long range inhibition.[6] Under those settings, the system represents a prototypic model of short term memory. When no stimulation is sent, the system remains quiescent (the activity is uniformly equal to a "resting" activity). When stimulated by the input signal at a given location, and for a certain time, the system displays a peak of activity in the vicinity of the stimulation, and this strong activity is stable and persistent even when the stimulus is withdrawn. This retention of dynamic activity relies on the same principle as that in in the Turing reaction-diffusion example, i.e. it corresponds to a stable equilibrium between excitation and inhibition trends. The location of the peak is not constrained (every map location is susceptible to develop a peak), and the number of simultaneous peaks at a given time is dependent on the range of inhibition. This system also displays a target tracking property, when the stimulus is moving in the field, and property of resetting when a stimulation occurs in a vicinity of an active peak, or when one sends a flash of activity in the whole field (external resetting). For their simplicity and robustness, such maps have a wide range of applications in robotics and control (Schöner et al., 1995; Bicho and Schöner, 1997; Moga and Gaussier, 1999).

Apart from engineering applications, an important interest has emerged in topologically organized models of short term memory in the cortex since the experiment of Funahashi et al. (Funahashi et al., 1989). Some studies have established the links between biology and neural map type models, interpreted as a mean-field approximation of neuronal activity (Ben-Yishai et al., 1995; Ben-Yishai et al., 1997), and pointed out some interesting properties relating to the adaptive sharpening where the input orientation is weakly contrasted. This model has also been implemented with biologically plausible neurons, where the same soft tuning property is established (Hansel and Sompolinsky, 1996). However, the property of retention is not explicitly established in those models. In parallel, other authors implemented a different kind of neural map as a model of short term memory, where retention properties rely on the use of bi-stable neurons (Camperi and Wang, 1998). This model has been extended to the use of conductance-type neurons with elaborate weight dynamics (Wang, 1999; Compte et al., 2000), where the stability of the memory properties rely on the changing of synaptic conductance during the retention process (gating mechanism with different time constants depending on the receptor).[7]

So despite numerous works on that subject, the question of dynamic retention in topologically organized structures is lacking on several points:

– There is still a lack of explanation about the exact contribution of local cellular dynamics on the global property of retention. One can thus ask whether global multi-stability is reachable without adaptive mechanism taking place at cellular or synaptic level.

– There is also a real opportunity in exploiting in artificial control systems the simplicity and robustness of neural maps retention properties, in particular through interactions with deeper neural structures and temporal events learning.

This paper will mainly focus on the first point with the presentation of a recurrent map of discrete binary spiking neurons, whereas the second consideration will be examined from the perspective of reinforcement learning and robot simulations.

3.1. *Modelling a neural map with a RRNN*

In order to model neural field properties, we introduce a topology in the weight structure. We define a parameter $r^{(pq)}$ which represents the neighbourhood density from population q towards population p (also called neighbourhood radius). The weights are thus adapted according to the (normalized) distance between neuron i and neuron j (we take a simple $1D$ closed ring geometry), i.e. $\forall i \in \{1, \ldots, N^{(p)}\}, \forall j \in \{1, \ldots, N^{(q)}\}$, we calculate the distance

$$\delta_{ij}^{(pq)} = 2\pi \times \min\left(\left|i/N^{(p)} - j/N^{(q)}\right|, 1 - \left|i/N^{(p)} - j/N^{(q)}\right|\right)$$

so that $\delta_{ij} \in [0, \pi]$, and then we calculate a gaussian normalized neighbourhood factor

$$\nu^{(pq)}(\delta) = \frac{\sqrt{2\pi}}{r^{(pq)}} \exp\left(-\frac{1}{2}\left(\delta/r^{(pq)}\right)^2\right)$$

We finally modify the weights according to this neighbourhood factor

$$\tilde{J}_{ij}^{(pq)} = J_{ij}^{(pq)} \times \nu^{(pq)}\left(\delta_{ij}^{(pq)}\right)$$

Knowing that

– $\int_0^\infty \nu^{(pq)}(\delta) d\delta = \pi$, and, for small $r^{(pq)}$, $\int_0^\pi \nu^{(pq)}(\delta) d\delta \simeq \pi$.
– The repartition of $\delta_{ij}^{(pq)}$'s is uniform in $[0, \pi]$,

For $r^{(pq)} < 1$, weights enhancements and weights decays are globally balanced, so that the expectation of the weights remains unchanged by this transformation. On the contrary, the weights standard deviation is increased by a factor of the order of $1/r^{(pq)}$. In order to avoid too large weights distortion, we define an adaptation factor which is designed in order to approach $1/r^{(pq)}$ for $r^{(pq)} < 1$, and 1 for $r^{(pq)} \geq 1$, i.e. $\kappa^{(pq)} = 1 + \frac{\exp\left(-\left(r^{(pq)}\right)^2\right)}{r^{(pq)}}$ and the eccentricity is adapted such that $d^{(pq)'} = \sqrt{\kappa^{(pq)}} d^{(pq)}$.

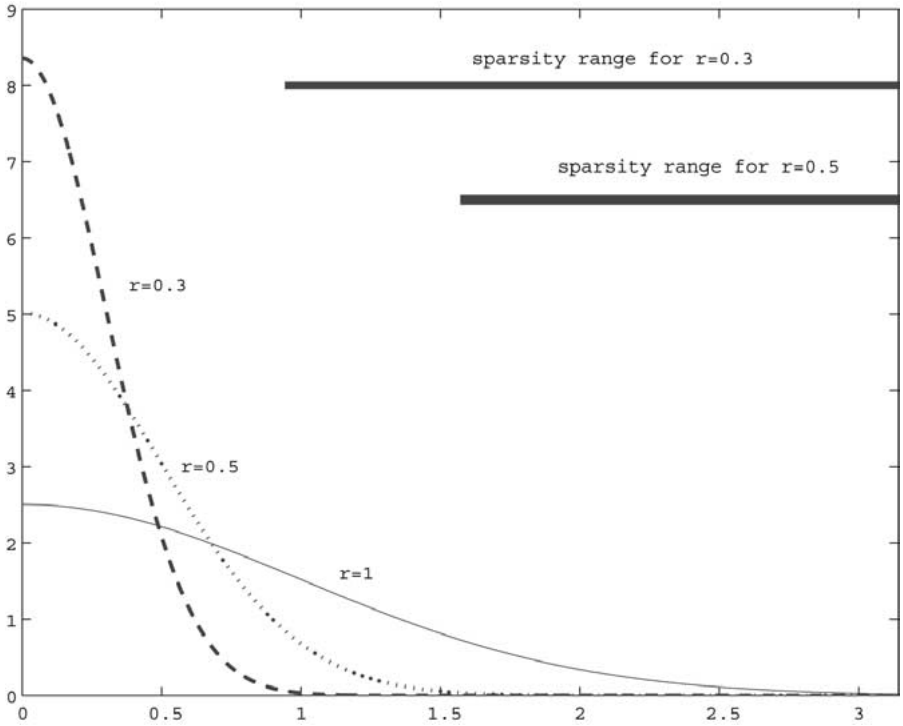

Figure 4. **Weights adaptation factor according to the distance**. The distance is on x-axis. The adaptation factor is given for $r = 1$ (solid line), $r = 0.5$ (dotted line), $r = 0.3$ (dashed line). For $r = 0.3$ and $r = 0.6$, we indicate the distance range for which the weights are set to zero.

The smaller the value of $r^{(pq)}$, the tighter the neighbourhood. Close links are strongly enhanced (according to the initial random draw) whereas distant links fade to zero (see Figure 4). One can notice that introducing a neighbourhood factor tends to increase the sparsity of the weights. In concrete terms, every link such that $d_{ij}^{(pq)} > \pi r^{(pq)}$ is suppressed, i.e. set to zero. It can be noticed that for $r^{(pq)} \leq 1$, the effective sparsity is $\rho^{(pq)'} = r^{(pq)}\rho^{(pq)}$.

3.2. *Simulation and neural map properties*

3.2.1. *Parameter setting*

In our simulations, parameters have been set such that:

- 75% of neurons are excitatory neurons, i.e. $N^{(1)} = 1000$ and $N^{(2)} = 300$.
- The excitatory delays remain faster than inhibitory delays. We take longer delays than in previous simulations, i.e.:

$$\tau_0 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad \lambda = \begin{pmatrix} 4 & 8 \\ 4 & 8 \end{pmatrix}$$

- The weight eccentricity is a bit stronger than in previous simulations, i.e. $k = 3$, $d = 4.5$.
- The activation thresholds $\theta^{(1)}$ and $\theta^{(2)}$ are set to 0.1.
- A field is defined on excitatory layer where the excitatory radius is set tighter than the inhibitory radius on excitatory layer, i.e. $r^{(11)} = 0.1$, $r^{(12)} = 0.3$. The range of excitation thus corresponds to $\simeq$10% of the map and the range of inhibition corresponds to $\simeq$30% of the map. No vicinity is defined for the inhibitory layer, so that any neuron of layer 1 may activate any neuron of layer 2.

Those settings define a discrete-time/binary states system which is functionally close to the continuous time/continuous state space of the neural field of eq.(6), where the continuous neighbour field integration is replaced by a weighted summation of incoming spike through delayed connections from excitatory and inhibitory layers. The homogeneous field of eq.(6) is replaced by inhomogeneous weights and delays, and the regularities in behaviour are to arise from the homogeneity of the probability distributions that define the weights (uniform distribution with neighbour reinforcement) and the delays (Poisson distribution).

3.2.2. *Simulations*

A typical system simulation is presented on Figure 5.

The simulation results are very robust for a change on the initial random setting of weights and delays, number of neurons in excitatory and inhibitory layers, etc. ... The parameters settings are not bounded to the given values, but the eccentricity of the weight parameters has to be stronger than 3, and inhibitory delays have to remain significantly longer than excitatory ones.

Under the given settings, the spontaneous dynamics of the system (when no particular input is sent) fades to zero, which is the quiet default dynamics of the system (this is due to the stronger eccentricity). Note however that the system is very close to the activation threshold, and thus very sensitive to any perturbation or excitation coming from outside (i.e. it has a natural tendency to amplify incoming perturbations).

Inputs are sent during short periods of time on the excitatory layer in order to test the reactivity and memory properties of the system. The first three stimuli (occurring at $t = 301$, $t = 501$ and $t = 701$) have globally the same shape (i.e. they locally excite small groups of neurons for a short time interval), but their duration (40 time steps for stim. 1 and 3, 20 time steps for stim. 2) and the number of neurons (10 neurons for stim. 1 and 3, 20 neurons

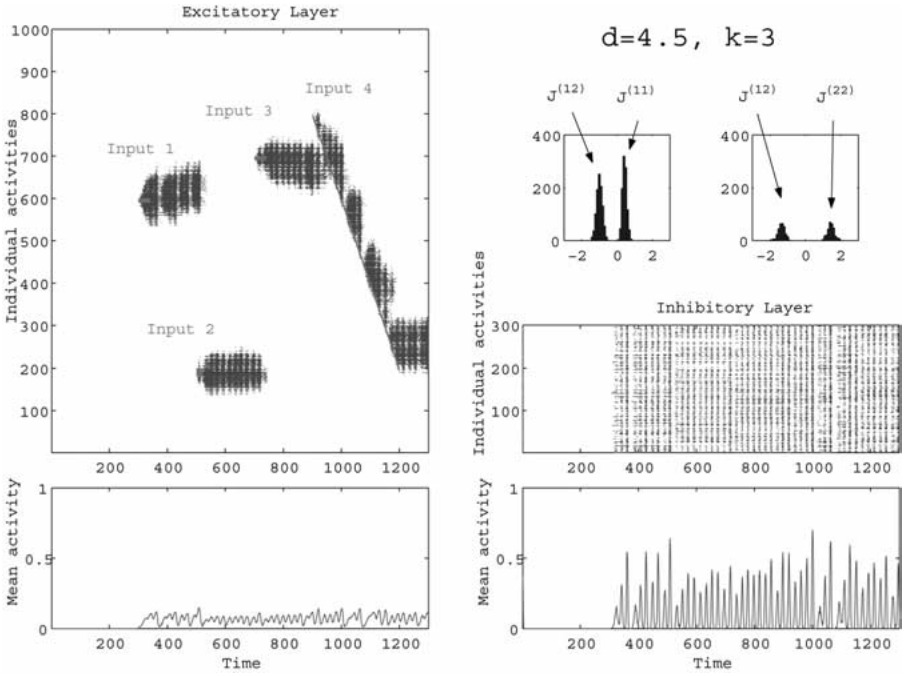

Figure 5. **Input presentation in a topologically organized network** (only the excitatory layer owns a topology). The network is composed of $N^{(1)} = 1000$ neurons and $N^{(2)} = 300$ neurons. The effective sparsity are $\rho^{(11)'} = 2.6\%$, $\rho^{(12)'} = 22.5\%$, $\rho^{(21)} = 7.6\%$, $\rho^{(22)} = 22\%$. Excitatory links are globally faster ($\simeq 5$ time steps) than inhibitory links ($\simeq 9$ time steps). Four inputs (red) are sent. The first input is sent at time 301–340 on neurons 591–600. The second input is sent at time 501–520 on neurons 181–200. The third input is sent at time 701–740 on neurons 691–700. The last input crosses the whole field, activating 10 neurons at each time step **Upper left**: global activity of the excitatory layer. **Lower left**: mean activity on excitatory layer. **Middle right**: global activity of the inhibitory layer. **Lower right**: mean activity on inhibitory layer. **Upper right insets**: distribution of the sum of incoming weights from excitatory and inhibitory populations, for 1000 excitatory neurons (left) and 200 inhibitory neurons (right).

for stim. 2) are not the same. Stimulus 4 is a ramp with a constant shift, lasting 300 time steps.

Figure 5 presents the activity of the excitatory layer during 1300 time steps, while those inputs are sent, with the parallel activity in inhibitory layer. The spontaneous activity, as we said, is quiescent (time 1–300) and neurons do not reach the activation threshold during this time interval.

Three main observations can be done:

– The activation rapidly propagates to the neighbourhood of the stimulated neurons. This propagation is stopped for long distances by the inhibitory influence which has been activated in parallel. The width of the activated

band in the excitatory layer is of the order of 80–100 neurons which corresponds to the neighbourhood rate $r^{(11)} = 0.1$.

- The neurons of inhibitory layer are fully synchronized with long waves of period $\simeq 20$ time steps. The activity is synchronous and bursting with a period of 20 approximately. This synchrony (in that case: synchronous chaos) is due to the flipping reverberant activity between layer 1 and layer 2, as mentioned in Figure 3.
- When the stimulus stops at time $t = 341$, the layer remains active and the dynamics is self-sustained, at the location which was previously stimulated. This persistent activity is a prototypic example of *dynamic retention*. This dynamic memory property gives rise to a fragile wave of activity which remains active until resetting caused by external influence.

It can be remarked that excitatory activity strongly fluctuates in time according to inhibitory synchronous oscillations. The global activity remains weak and fragile, so that any input, even small and short, leads to a strong modification of the neuronal activity. *The property of retention does not overtake the property of adaptivity*.

Figure 5 gives a global view on the activity caused by the stimuli. Stimuli 2 and 3 give rise to waves of activity which do not differ in shape, period and radius from the first wave. Even stim. 2 which strongly differs from stim. 1 in range and duration gives rise to a comparable wave of activity. Our system thus realizes a *normalization* of the input signal.

At last, one can observe a tracking behaviour when stim. 4 is presented. The wave of activity follows the input until the stimulus is taken off, and then remains straight. The tracking property is thus bounded to persistent stimuli, as the system does not interpolate the speed of the target.

4. Discussion

We have presented here simulations results on a large size random recurrent neural network model with simple discrete-time spiking activation function. In a model with two populations, one of "interactive" excitatory neurons, one of inhibitory "interneurons", and taking inspiration from Amari's neural field principles, we show that one can obtain robust *retention* and *normalization* of incoming stimuli. These results, which extend Amari's model to the spiking neurons domain, makes it more plausible and may bring new insights in the modelling of topologically structured brain maps.

Contrary to previous simulation work (Bush and Sejnowski, 1996), we suggest here that inhomogeneous delays may favour synchronization in excitatory-inhibitory networks, provided that inhibitory delays distribution

is significantly longer than the excitatory one. In that case, slow waves of synchronous activity appear (this synchronous activity is stronger in inhibitory population). This synchrony allows for a fast adaptivity of the network to various input stimuli. The temporal scales we used suggest to identify those slow waves to the gamma frequency activity (30–80 Hz) which is very commonly observed in cortical dynamics. The rates of spiking activity (10% of activity), the sparsity of weights and delays, the strong adaptivity of the dynamics that are displayed by the model helps to approach a better understanding of the real dynamic properties of biological neural networks. However, at the present time, our system doesn't give clues on the functional activity that may underlie synchronous activation.

An ongoing project in the "Movement and Perception" lab is to apply this system to the modelling of the superior colliculus maps, and to the adaptive properties taking place in this structure which is the core of basic visuo-motor transformations (Guillaume and Pélisson, 2001) (see also (Trappenberg, 1998) for colluculus-inspired simulation work). Another possible application could be the embedding of this network in a robotic adaptive control structure, as its spontaneous properties are of interest for intermittent target tracking in a noisy environment where targets are of various sizes, duration and intensity. The question of learning and adaptation is still under investigation in our networks. As in (Daucé et al., 2002), the presented system is supposed to be able to learn temporal sequences with a simple Hebbian learning rule. At last, one can ask the question of giving a solid theoretical support for analyzing the regimes of those topological structured random maps at the limit of large sizes, which is, to our knowledge, not yet established at the present time.

Acknowledgements

This work is partially supported by ACI "neurosciences intégratives et computationnelles", thème "temps et cerveau". Thanks to Manuel Samuelides, Bruno Cessac and Laurent Perrinet for fruitful discussions.

Notes

[1] On the contrary to Hopfield system (Hopfield, 1998), the input doesn't correspond to the initial state $x_i^{(p)}(0)$ of the network.

[2] It is a bit surprising that the question of transmission delay is so rarely tackled in the field of spiking neural networks. The emphasis is often put on the spike emission mechanisms and coincidence detection whereas transmission delays are often considered as uniformly constant from one layer to the other (Thorpe et al., 1996).

[3] With variance $\rho^{(pr)}(1-\rho^{(pr)})N^{(r)} = (1-\rho^{(pr)})N_{\max}^{(pr)} = \left(1 - \frac{N_{\max}^{(pr)}}{N^{(r)}}\right) N_{\max}^{(pr)} \simeq N_{\max}^{(pr)}$ for large $N^{(r)}$.

[4] More generally, we set $\tau_{\max}$ initial vectors corresponding to $\mathbf{x}(0), \mathbf{x}(-1), \ldots, \mathbf{x}(1-\tau_{\max})$ where $\tau_{\max}$ is the maximum delay in the system.

[5] This system is auto-catalytic at short range, and antagonist at long range (i.e. reaction catalysts diffuse with a slow speed, and counter-reaction catalysts diffuse with a higher speed). An equilibrium is obtained when the reaction compounds form a pattern of uniformly distributed aggregates, with a certain spatial period.

[6] One can note that Mexican hat kernels have a wide range of application in Self organizing Maps (Kohonen, 1982), but self organized classification systems are not in the scope of this paper.

[7] One can remark that other models of short term memory have been proposed, which do not rely on the topology of neural maps, e.g. (Zipser, 1993). The most well known is Amit and Brunel's model (Amit and Brunel, 1997), which use modular architectures with locally exciting connections and globally inhibiting connections (which is by some respect a roughly discretized neural map, with an identical principle of local excitation an global inhibition).

References

Amari S (1977) Dynamics of pattern formation in lateral-inhibition type neural fields. Biological Cybernetics 27: 77–87

Amit DJ and Brunel N (1997) Global spontaneous activity and learned local delay activity in cortical conditions. Cerebral Cortex, 7, 2 (1997) 7(2): 237–252

Ben-Yishai R, Hansel D and Sompolinsky H (1997) Traveling waves and the processing of weakly tuned inputs in a cortical network model. J. Comp. Neurosci. 4: 57–79

Ben-Yishai R, Lev Bar-Or R and Sompolinsky H (1995) Theory of orientation tuning in visual cortex. Proc. Nat. Acad. Sci. USA 92: 3844–3848

Bicho E and Schöner G (1997) The dynamics approach to autonomous robotics demonstrated on a low-level vehicle platform. Robotics and Autonomous System 21: 23–35

Bush P and Sejnowski T (1996) Inhibition synchronizes sparsely connected cortical neurons within and between columns in realistic network models. J. Comput. Neurosci. 3(2): 91–110

Camperi M and Wang X-J (1998) A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. J. of Computational Neuroscience 5: 383–405

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

Compte A, Brunel N, Goldman-Rakic PS and Wang X-J (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. Cerebral Cortex 10: 910–923

Daucé E, Quoy M and Doyon B (2002) Resonant spatiotemporal learning in large random recurrent networks. Biological Cybernetics 87: 185–198

Doyon B, Cessac B, Quoy M and Samuelides M (1993) Control of the transition to chaos in neural networks with random connectivity. Int. J. of Bif. and Chaos 3(2): 279–291

Funahashi S, Bruce CJ and Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J. Neurophysiol. 61: 331–349

Gerstner W (1999) Spiking Neurons, Chapt. 1, pp. 3–33. MIT Press, Cambridge, MA.

Guillaume A and Pélisson D (2001) Gaze shifts evoked by electrical stimulation of the superior colliculus in the head-unrestrained cat. I. Effect of the locus and of the parameters of stimulation. European Journal of Neurosciences 14: 1331–1344

Hansel D and Sompolinsky H (1996) Chaos and synchrony in a model of a hypercolumn in visual cortex. J. Comp. Neurosc. 3: 7–34

Hermann M, Hertz J and Prugel-Bennett A (1995) Analysis of synfire chains. Network 63(3): 403–414

Herz A, Sulzer B, Kuhn R and van Hemmen JL (1989) Hebbian learning reconsidered: representation of static and dynamic objects in associative neural nets. Biol. Cybern. 60: 457–467

Hopfield J (1998) Neural networks and physical systems with emergent collective computational abilities. Biological Cybernetics 79: 109–120

Hoshino O, Kashimori Y and Kambara T (1982) An olfactory recognition model based on spatio-temporal encoding of odor quality in the olfactory bulb. Proc. Nat. Acad. Sci. 79: 2554–2558

Hupé J-M, James A, Payne B, Lomber S, Girard P and Bullier J (1998) Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. Nature 394: 784–787

Izhikevich EM (2003) Simple model of spiking neurons. IEEE trans. on neural networks 14(6): 1569–1572

Kohonen T (1982) Self-organized formation of topologically correct feature maps. Biological Cybernetics 43: 59–69

Mac Cullogh WS and Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5: 115–133

Mac Leod K and Laurent G (1996) Distinct mechanisms for synchronization and temporal patterning of odor-encoding cell assemblies. Science 274: 976–979

Moga S and Gaussier P (1999) A neuronal structure for learning by imitation. In: Floreano D, Nicoud J-D and Mondada F (eds) Lecture Notes in Artificial Intelligence – European Conference on Artificial Life, pp. 314–318 Lausanne.

Moynot O and Samuelides M (2002) Large deviations and mean-field theory for asymmetric random recurrent neural networks. PTRF (123-1): 41–75

Nowak L and Bullier J (1997) The timing of information transfer in the visual system. In: R. et al. (ed) Cerebral Cortex, pp. 205–240. Plenum Press, New York.

Nützel K, Kien J, Bauer K, Altman J and Krey U (1994) Dynamics of diluted attractor neural networks with delays. Biol. Cybern. 70: 553–561

Riedel U, Kuhn R and Van Hemmen J (1988) Temporal sequences and chaos in neural nets. Physical review A 38(2): 1105–1108

Schöner G, Dose M and Engels C (1995) Dynamics of behavior: theory and applications for autonomous robot architectures. Robotics and Autonomous System 16(2–4): 213–245

Shannon C and Weaver W (1949) The Mathematical Theory of Communications. University of Illinois Press, Chicago

Skarda C and Freeman W (1987) How brains make chaos in order to make sense of the world. Behav. Brain Sci. 10: 161–195

Sompolinsky H, Crisanti A and Sommers H (1988) Chaos in random neural networks. Phys. Rev. Lett. 61: 259–262

Thorpe SJ, Fize D and Marlot C (1996) Speed of processing in the human visual system. Nature 381: 520–522

Trappenberg T (1998) Dynamic cooperation and competition in a network of spiking neurons. In: Proceedings of the International Conference on Neural and Intelligent Processing ICONIP'98

Tsodyks MV and Feigelman (1988) Enhanced storage capacity in neural networks with low level of activity. Europhysics Letters 6: 101–105

Turing A (1952) The chemical basis of morphogenesis. Philos. Trans. R. Soc. London: 37–72

Wang X-J (1999) Synaptic basis of cortical persistent activity: the importance of NMDA receptors to working memory. The Journal of Neuroscience 19(21): 9587–9603

Zipser D (1993) Recurrent network model of the neural mechanism of short-term active memory. Neural Computation 3: 179–193