Dynamics of Large Random Recurrent Neural Networks : Oscillations of 2-Population Model

MOYNOT Olivier, SAMUELIDES Manuel
ONERA-CERT (DTIM), 2 avenue Edouard Belin, 31055 Toulouse cedex.
e-mails : Olivier.Moynot@cert.fr, Manuel.Samuelides@cert.fr

Abstract

We are interested in the asymptotic behavior of noisy discrete time neural networks with two populations of neurons. The couplings and thresholds are asymmetric and gaussian. We use the large deviation techniques developed by Ben Arous and Guionnet to study the limit behavior of our networks when their size grows to infinity. We prove a propagation of chaos property, which is closely related to vanishing correlations of activation states. We are also able to compute the limit distribution of the activation potentials of the neurons in the thermodynamic limit. It is gaussian and characterized by a set of dynamic mean-field equations. The numerical study of these equations reveals a parametric domain where the mean of this limit law is subject to periodic oscillations. This property can be directly related to synchronization. Moreover, we prove a useful equation satisfied by the mean quadratic distance between two trajectories, which allows to predict the dynamics of the network.

Introduction

Large random networks and their relations to particle systems and especially spin glasses has interested many biologists, physicists and mathematicians for two decades. The major goal of these investigations is to obtain the distribution of the activation potential of the neurons when the size of the network grows to infinity. The characteristics of this limit law are embedded in the mean-field equations.

After Amari's first works (see [1]), the interest focused on networks with asymmetric couplings, which variance is of order 1/N, where N is the number of neurons. Geman [8] proved the convergence property in some particular cases, and notably for linear models. Then Sompolinsky (see [5,14]) used statistical physics methods to obtain these mean-field equations for gaussian connection weights and to study the dynamical properties of the associated continuous time networks. Cessac *and al.* [3,4] used the same approach for discrete time models and numerically showed the general occurrence of chaos by a quasi-periodicity route in large size networks. Moreover, they established that this asymptotic regime is described in the thermodynamic limit by the mean-field equations. They also obtained the vanishingcorrelations of activation states in their fully connected neural networks.

The first purpose of this paper is to extend these results to our 2-population model, and to give a rigorous proof for them. We use extensively the ideas developed in [2,9] by Ben Arous and Guionnet. In these papers, they developed large deviation techniques to establish many properties in a continuous time spin glasses context. They considered one population of spins, and their couplings were gaussian and centered. They proved the weak convergence of the law of every spin towards a measure given by an implicit equation. They also obtained a propagation of chaos result. Although the deep relations between our two populations and the particularities of our model slightly increase the complexity of the proof, we use the same methods to deduce the propagation of chaos and the mean-field equations. The whole rigorous demonstration can be found in [6]. Moreover, these large deviation techniques lead to an equation which describes the evolution of the mean quadratic distance between two initially independent trajectories.

The second part of this communication is dedicated to the dynamical behavior of our networks. We obtain properties in the thermodynamic limit which are in good agreement with the results of numerical simulations of large size recurrent networks. This notably means that the mean-field equations and the mean quadratic distance are of great help to anticipate the dynamical properties of our large finite size models.

More precisely, numerical computations realized on these equations reveal a parametric domain where the neurons get synchronized for large time. This is related to many recent biological discoveries (see [10,12] for example), which underline the great importance of synchronization in neural dynamics in the brain. Notice that such a behavior

0-7803-5529-6/99/$10.00

doesn't occur in the single population model studied by Cessac. Furthermore, the expression proved for the mean quadratic distance between two trajectories gives a criterion to characterize the occurrence of chaos in our networks.
Collecting these properties together, we obtain the bifurcation map of the network for some particular values of the parameters and give prominence to different dynamic regimes.

The model

We consider the following discrete time recurrent neural network, with dynamics :

$$x_i^p(t) = f(u_i^p(t))$$

$$u_i^p(t) = \sum_{j=1}^{n_1} J_{ij}^{p1} x_j^1(t-1) + \sum_{j=1}^{n_2} J_{ij}^{p2} x_j^2(t-1) + \sigma W_i^p(t) - \theta_i^p$$

Our network contains two populations of neurons, whose number is given by exponent p. They might for instance represent excitators (p=1) and inhibitors (p=2). There are n_p neurons of population p. For $p,q \in \{1,2\}^2$, the (J_{ij}^{pq})'s represent the connection weights relative to the influence of population q on population p. The (θ_i^p)'s are the thresholds, and $W_i^p(t)$ is a synaptic noise. f is an arbitrary sigmoid function taking its values into $]0,1[$ (for all the concrete applications, we take $f(x)=(1+th(gx))/2$). $x_i^p(t)$ represents the activation state of the neuron i of population p at time t. It corresponds to the spikes discharge frequency of the neuron. All the neuron's activation states are supposed to be independent at time 0 with respective initial laws μ_0^1 and μ_0^2 for the two populations. $u_i^p(t)$ is the activation potential of the neuron i of population p at time t.
Our nets are fully connected. We suppose that the distributions of the connection weights, the thresholds and the synaptic noise are respectively gaussian laws $N(\frac{\bar{J}^{pq}}{n_q}, \frac{(J^{pq})^2}{n_q})$, $N(\bar{\theta}^p, (\theta^p)^2)$ and $N(0,1)$. All these random variables are supposed to be independent. We study the evolution of this system when the sizes of the populations grow to infinity without any change in their proportion.

Mathematical advances

We consider the evolution of the system between 0 and a fixed time T. For technical reasons, we suppose $\sigma > 0$. Notice that the results are valid for an arbitrary small noise, and that simulations confirm their generality. Let Q^N represent the law of all the activation potentials of the network. We then use our large deviations techniques and the ideas developed in [13] (all the activation states of the neurons of a given population have the same distribution).
We deduce the following **propagation of chaos result :**
Let k_1, k_2 be two integers, and $h_1^1,\ldots,h_{k_1}^1, h_1^2,\ldots,h_{k_2}^2$ be a set of bounded continuous test functions taking real values. Then there are two probabilities Q^1 and Q^2, defined on $]0,1[^{\{0,\ldots,T\}}$ and gaussian for t greater than 1, such that :

$$\int \prod_{p=1}^{2} \prod_{i=1}^{k_p} h_i^p(u_i^p) dQ^N(u) \xrightarrow[N\to+\infty]{} \prod_{p=1}^{2} \prod_{i=1}^{k_p} \int h_i^p(u_i^p) dQ^p(u_i^p)$$

This propagation of chaos result is a mathematical strong statement corresponding to the vanishing correlations hypothesis of the physicists. Let us now explain it in concrete terms : at time 0, the activation states of the neurons are chosen independent from each other. But from time 1 to T, as the net in fully connected, many relations take place between the neurons. We call propagation of chaos the property of the activation potentials u_i^p to behave asymptotically as *independent* random vectors when the size N of the network grows to infinity.
This particularly implies that all the neurons of every population tend to behave as a generic asymptotic neuron, whose activation potential's law is Q^p.

In our discrete time context, we are able to compute the gaussian characteristics of Q^p. We consider $(\mu^p(t), \Delta^p(t,t'))_{1\le t,t'\le T}$ the expectation and covariance matrix of Q^p. In particular, we note $v^p(t) = \Delta^p(t,t')$. Δ^p represents the time covariance of each population's generic neuron. We also note :

$$m^p(t) = \int f(u_t) dQ^p$$

$$q^p(t) = \int f^2(u_t) dQ^p$$

Finally, let $Dh = 1/\sqrt{2\pi}\exp(-h^2/2)$.
We then have the following **mean-field equations :**
For $1 \le t,t' \le T$ and $t \neq t'$,

$$m^p(t) = \int f(\sqrt{v^p(t)}h + \mu^p(t)) Dh$$

$$q^p(t) = \int f^2(\sqrt{v^p(t)}h + \mu^p(t)) Dh$$

$$\mu^p(t+1) = -\bar{\theta}^p + \bar{J}^{p1} m^1(t) + \bar{J}^{p2} m^2(t)$$

$$v^p(t+1) = (\sigma)^2 + (\theta^p)^2 + (J^{p1})^2 q^1(t) + (J^{p2})^2 q^2(t)$$

$$\Delta^p(t+1,t'+1) = (J^{p1})^2 C^1(t,t') + (J^{p2})^2 C^2(t,t') + (\theta^p)^2$$

$$C^p(t,t') = \int DhDh' f(a^p(t,t')) f\left(h'\sqrt{v^p(t')} + \mu^p(t')\right)$$

$$a^p(t,t') = \frac{\sqrt{v^p(t)v^p(t') - (\Delta^p(t,t'))^2}}{\sqrt{v^p(t')}} h + \frac{\Delta^p(t,t')}{\sqrt{v^p(t')}} h' + \mu^p(t)$$

As they depend on a small set of parameters (in particular they don't depend on the size N of the network), these equations are of great help for anticipating the dynamics of large neuronal assemblies (see next section).
Furthermore, the exponentially fast convergence properties associated to the large deviations principle lead to the following **law of large numbers** :

For any integer N, let $(u^1_{k,N}, u^2_{m,N})_{1\le k\le n_1, 1\le m\le n_2}$ be a family of random variables with law Q^N. Then for any $p\in \{1,2\}$, almost surely,

$$\frac{1}{n_p}\left(\sum_{i=1}^{n_p} f(u^p_{i,N}(t))\right)\xrightarrow[N\to+\infty]{}\int f(u_t)dQ^p = m^p(t)$$

This theorem gives a convergence result for almost all the choices of the parameters of the networks. This property allows us to use the mean-field equations to predict the behavior of macroscopic observable of a particular instantiation of the network.

Moreover, the large deviations techniques we use give access to the study of the **mean-quadratic distance** between two given trajectories.

For any $p\in \{1,2\}$, we consider $u^p(t)=(u^p_i(t))_{1\le i\le n_p}$ and $v^p(t)=(v^p_i(t))_{1\le i\le n_p}$, with following dynamics :

$$u^p_i(t)=\sum_{j=1}^{n_1} J^{p1}_{ij} f(u^1_j(t-1))+\sum_{j=1}^{n_2} J^{p2}_{ij} f(u^2_j(t-1)) + \sigma W_i^p(t)-\theta^p_i$$

$$v^p_i(t)=\sum_{j=1}^{n_1} J^{p1}_{ij} f(v^1_j(t-1))+\sum_{j=1}^{n_2} J^{p2}_{ij} f(v^2_j(t-1)) + \sigma \overline{W}_i^p(t)-\theta^p_i$$

The distributions of $u^p_i(0)$ and $v^p_i(0)$ are identical, and these two random variables are independent from each other. Notice that the parameters of the network are the same for u^p and v^p, except the noises which are supposed to be independent. As s can be chosen as small as we want, we suppose that the following study remains true without any noise.

We already know that the distributions of u^p_i and v^p_i converge separately towards the same law Q^p. We use our large deviations methods to study the covariance between u^p_i and v^p_i. We denote by R^N the global law of (u^p_i, v^p_i). The mean quadratic distance between u^p and v^p is defined as in [7] :

$$(d^p(t))^2=\lim_{N\to+\infty}\frac{1}{n_p}\sum_{i=1}^{n_p}\int [u^p_i(t)-v^p_i(t)]^2 dR^N$$

We prove that this mean quadratic distance satisfies the following relations :

$$(d^p(t))^2 = 2(v^p(t)-\Delta^p(t))$$

where :

$$\Delta^p(1)=\left(\int x_0 d\mu^p_o\right)^2$$

$$\Delta^p(t+1)=(J^{p1})^2 C^1(t)+(J^{p2})^2 C^2(t)+(\theta^p)^2$$

$$C^p(t)=\int DhDh' f(a^p(t)) f\left(h'\sqrt{v^p(t)}+\mu^p(t)\right)$$

$$a^p(t)=\frac{\sqrt{(v^p(t))^2-(\Delta^p(t))^2}}{\sqrt{v^p(t)}}h+\frac{\Delta^p(t)}{\sqrt{v^p(t)}}h'+\mu^p(t)$$

The notations are voluntarily chosen to underline the similarities of these equations with the mean-field equations.
This result is of great help to understand the behavior of our networks : the evolution of this mean quadratic distance for close initial conditions allows to know whether the dynamic regime of the system is chaotic (see [4,7] for instance).

Dynamic properties

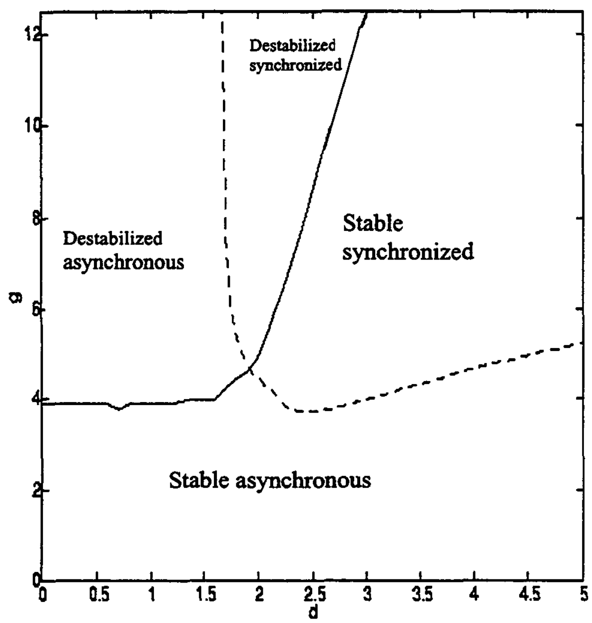
This section is dedicated to study the equations obtained before numerically and to give prominence to the various dynamic regimes of the network. We consider two populations of neurons, respectively composed of inhibitors and excitators. In seek of simplicity, we reduce the number of parameters and suppose that there is no inhibition on the inhibitors. More precisely, we suppose $J^{11}=J^{21}=J, J^{22}=0, J^{12}=\sqrt{2}J, \overline{J}^{11}=\overline{J}^{21}=Jd, \overline{J}^{22}=0$ and $\overline{J}^{12}=-2Jd$. Notice that d represents the "strength" of inhibition and excitation. In particular, for $d=0$, there is no qualitative difference between the neurons of the two populations.
The studies realized on the mean-field equations reveal that the average observable $m^1(t)$ and $m^2(t)$ are subject to periodic oscillations for certain values of the parameters. We say that our system is synchronized as soon as the signals $m^p(t)$ are not static for long time. This definition, which has already been used in [11], implies that the mean activity of the different neurons of a given population tends to evolve the same way.
In the single population model, Cessac established in [4] that the behavior of the neurons could be described by a stationary process in the thermodynamic limit (its mean and variance are constant temporal functions). In the two population model, the limit process associated to the mean-field equations is cyclostationary (its mean and covariance matrix are periodic temporal functions).
Remark here that there is no contradiction for the neurons to be synchronized and independent in the same time : the synchronization is directly related to the temporal oscillations of $m^p(t)$, while the independence corresponds

to random individual fluctuations of the activation states around this mean.

The second dynamic characteristic of our networks we want to underline is expressed by the mean quadratic distance between two trajectories. We consider two close initial conditions. We use the mathematical results obtained in the previous section and numerical studies to compute the temporal evolution of $d^p(t)$. We say that our dynamics are *stable* if the mean quadratic distance $d^p(t)$ converges towards zero when t grows to infinity. We talk about *destabilized dynamics* if $d^p(t)$ remains large in the same conditions. Such a behavior is directly connected to chaotic properties of our networks. Remark here that the evolution of $d^p(t)$ presents the same qualitative characteristics as in the single population model.

Bifurcation map

We give here the bifurcation map obtained for our network.



We recall here that g is double of the gain parameter associated to the sigmoid function $f(x)=1/2(1+th(gx))$ and that $d = \frac{\bar{J}^{11}}{J^{11}}$ represents the intensity of inhibition and excitation.

We obtain four dynamical regions, delimited by two frontiers : the continuous line gives the destabilization of the mean-field process, and is deduced by studying the mean quadratic distance. The dashed line corresponds to the transition between asynchronous dynamics ($m^p(t)$ converges towards a fixed point) and synchronized dynamics ($m^p(t)$ oscillates for large time).

An important property of this bifurcation map is that synchronization occurs only if d is large enough : this underlines the paramount importance of the intensity of inhibition and excitation in the synchronization of the neurons. The effective presence of two well-separated populations in the network is necessarily to obtain these phenomena.

Conclusion

The mathematical and numerical work realized in this paper proves that the 2-population model presents some synchronization properties which don't take place in the single population model. This property is linked to periodic oscillations of $m^p(t)$ in the mean-field equations. Moreover, simulations executed recently establish that the behavior of $m^p(t)$ can even be chaotic in another range of parameters. This dynamical diversity confirms the interest of the 2-population model.

References

[1] Amari S. (1972). *Characteristics of random Nets of analog Neuron-like Elements,* IEEE Trans. Syst. Man. Cyb. Vol SMC-2 N5.

[2] G. Ben Arous, A Guionnet (1995). *Large deviations for langevin spin glass dynamics,* PTRF 102, 455-509.

[3] Cessac B., Doyon B., Quoy M., Samuelides M. (1994). *Mean-Field equations, bifurcation map and route to chaos in discrete time neural networks,* Physica D 74 24,44.

[4] Cessac B (1995). *Increase in complexity in random neural networks.* Physical journal I, 5 : 409-432.

[5] A.Crisanti, HJ.Sommers and H.Sompolinsky (1990). *Chaos in neural networks : chaotic solutions.* Preprint.

[6] Dauce, E., Moynot, O., Pinaud, O., Samuelides M. (1999). *Etude de la dynamique de réseaux neuronaux aléatoires récurrents,* DRET technical repart RT3/3739.00/DTIM, Onera.

[7] Derida B. and Pommeau Y. (1986). Europhysical letters, 1 : 45-59.

[8] Geman S., Hwang CR. (1982). *A chaos hypothesis for some Large Systems of random equations,* Z. Wahr. Verw. Gebiete 60, 291-314.

[9] A.Guionnet (1997). *Averaged and quenched propagation of chaos for spin glass dynamics.* PTRF, 109:183-215.

[10] Gray, C. M., Konig, P., Engel, A.K. and Singer, W. (1989). *Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflect global stimulus properties*, Nature 7, 338-334.

[11] Hansel, D. and Sompolinsky, H. (1996). *Chaos and synchrony in a model of a hypercolumn in visual cortex*, J. Comp. Neurosc. 3, 7-34.

[12] MacLeod, K., and Laurent. G. (1996) *Distinct mechanisms for synchronization and temporal patterning of odor-encoding cell assemblies*, Science 274, 976-979.

[13] Sznitman AS. (1984). *Equations de Boltzmann spatialement homogenes.* Z.F.Wahrscheinlichkeitstheorie Verw. Gebiete. 66 : 550-592.

[14] Sompolinsky H., Crisanti A., Sommers HJ. (1988). *Chaos in random neural networks,* Phys.Rev.Lett. 61 259-262.