Stochastic analysis of the dynamics of a general class of synchronous neural networks

1993 J. Phys. A: Math. Gen. 26 6879

(http://iopscience.iop.org/0305-4470/26/23/030)

J. Phys. A: Math. Gen. 26 (1993) 6879–6892. Printed in the UK

Stochastic analysis of the dynamic of a general class of synchronous neural networks

O Bernier

LAB/OCM/TEP, Centre National d'Étude des Télécommunications Route de Trégastel, 22301 Lannion Cedex, France

Received 14 June 1993, in final form 23 August 1993

Abstract. We present a generating-function-based method to analyse the dynamics of synchronous neural networks. It is applied to a very general model of networks of automata, which includes a certain number of neural network models of the Little type as particular cases. The method enables us to analyse the evolution of the order parameters as well as their fluctuations. These fluctuations are shown to consist of two independent parts, one due to the stochastic nature of the evolution, the other to the random nature of the stored patterns.

1. Introduction

In this article, we present a generating-function-based method to analyse the stochastic evolution of synchronous neural networks, such as the Little model [1, 2], in the case of low loading and at non-zero temperature. The dynamic of the order parameters has already been solved even for asymmetrical connections for this particular case [3]. Here we extend these results to a very general network model, where we tried to impose the minimum set of hypotheses compatible with the method. This model includes the case of the simple asymmetrical Little model already solved, but also various generalizations, such as the nonlinear Little model, or the extension of -1, $+1$ states to Potts states for the neurons ([4] and [5], for example, describe such models of neural networks and give results for the asynchronous case or for the one-step dynamic at saturation). Moreover, we derive not only the evolution of the global order parameters, but also their fluctuations up to first order in the system size expansion. Such a fluctuation analysis has been performed, for an asynchronous equivalent of the asymmetrical Little model by Shiino [6]. But to our knowledge, this has not been done for synchronous networks. By our method we are able to study the relative role of the two sources of fluctuations, the stochastic nature of the evolution, and the random nature of the stored patterns. We are also able to clarify the relative importance of the various assumptions of the model for the validity of the final results.

2. Description of the general model

The model we will use is a very general network model of parallel synchronous stochastic automata. We consider a network of N automata, indexed by i, each one with a state σ_i. The set of possible values of the state is not important, as long as there exist a measure on it, which will provide a mean to define a probability. To each automaton we also associate an element ξ_i, belonging to a finite set. It can be thought of as a kind of local memory for the automaton, and corresponds in the special case of the Little model to the stored patterns. We will consider the case where for each i, ξ_i is chosen randomly, independently of the others, with an identical probability for each automaton denoted by $k(\xi)$. All the automata evolve in parallel, synchronously: at each time step t, each one updates its state independently of the others. The probability governing this update is supposed to be defined by a positive bounded function $\mathbf{T}$, function of the value of ξ for this automaton and of a certain number of global parameters of the network denoted by m_ρ (ρ being a discrete index, with r possible values) and given by

$$m_\rho = \frac{1}{N} \sum_{i=1}^{N} M^\rho(\sigma_i, \xi_i) \tag{2.1}$$

where the M^ρ are r functions of the state of an automaton and of its 'memory', which are supposed to be bounded, so that the parameters are bounded when N grows. These parameters correspond to the mean value for all neurons of these functions. $\mathbf{T}$ must be derivable in terms of the global parameters. This gives the following probability of finding the automaton indexed by i in the state σ_i^{t+1} at time $t+1$:

$$P(\sigma_i^{t+1}) = \frac{\mathbf{T}(\sigma_i^{t+1}, \xi_i, \{m_\rho^t\}_\rho)}{\displaystyle\int \mathrm{d}\sigma\, \mathbf{T}(\sigma, \xi_i, \{m_\rho^t\}_\rho)}. \tag{2.2}$$

Here $\{m_\rho^t\}_\rho$ denotes all the parameters m_ρ^t with t fixed and for all values of ρ. The measure defined on the set of the values of σ is supposed to be well behaved, so that all integrals involving measurable bounded functions are convergent. (For example, the classical integration over a compact interval of $\mathbb{R}$ verifies this.) Moreover, the integral defining the denominator is supposed to be non-zero for every values of the parameters ξ and of the m_ρ^t. We analyse in this article the case of N growing to infinity with r and the number of possible values of ξ fixed. We will use ρ, ρ^*, or η, η^*, κ to index the r global parameters, while t, t', τ, τ', u will be used to index a particular time. The evolution will be studied between time 0 and a maximum time T.

To illustrate the applicability of this model, we describe how the asymmetrical Little model can be considered a special case of it. It corresponds to the choice $\{-1, +1\}$ as the set of possible values of the state of the automaton (in this case called neuron). The ξ_i belongs then to $\{-1, +1\}^p$ and for each μ between 1 and p, ξ_i^μ corresponds to the value of the neuron i in the μth stored pattern. In this particular case, the index ρ is identified with μ. The global parameters are the correlations between the states of the network as a whole and the stored patterns:

$$m_\mu = \frac{1}{N} \sum_{i=1}^{N} \sigma_i \xi_i^\mu. \tag{2.3}$$

The function governing the stochastic evolution of the network is given by

$$\mathbf{T}(\sigma_i^{t+1}, \xi_i, \{m_\mu^t\}_\mu) = \exp \beta \left\{ \sigma_i^{t+1} \sum_{\mu\nu} \xi_i^\mu A_{\mu\nu} m_\nu^t + \sigma_i^{t+1} \sum_\mu \xi_i^\mu \Theta_\mu \right\}. \tag{2.4}$$

This corresponds to a Little model with temperature parameter β and with the connections and thresholds given by

$$J_{ij}=\frac{1}{N}\sum_{\mu\nu} A_{\mu\nu}\xi_i^\mu\xi_j^\nu \qquad \Theta_i=\sum_\mu \Theta_\mu\xi_i^\mu \tag{2.5}$$

where the ξ_i^μ are the p-stored patterns, and where A is not constrained to be symmetrical [3].

3. Analysis of the evolution

To analyse the stochastic evolution of this general type of network, we introduce the following generating function, which gives by derivation all the moments of the repartition of the parameters m_ρ^t, for all values of ρ and t (t between 0 and T):

$$\Gamma(\{\lambda_\rho^t\}_{\rho t})=\int\prod_{it}\mathrm{d}\sigma_i^t\prod_{t\neq T}\prod_i\frac{\mathbf{T}(\sigma_i^{t+1},\xi_i,\{m_\rho^t\}_\rho)}{\int \mathrm{d}\sigma\mathbf{T}(\sigma,\xi_i,\{m_\rho^t\}_\rho)}\exp F(N,\{m_\rho^0\}_\rho)\exp\sum_{\rho t}\lambda_\rho^t m_\rho^t \tag{3.1}$$

where F is the initial probability of the state of the network, that we suppose to depend only on the m_ρ^0 and N. m_ρ^t are the global parameters defined as functions of the σ_i^t by (2.1). They are also, as (2.2) shows, the important parameters of the network (they are sufficient to give the probability of each state of the network at the next time step when the ξ_i are given). We can insert in this equation a delta function for each relation corresponding to (2.1), at each time t, and as a consequence introduce the m_ρ^t as independent integration variables. Using the integral representation of the delta function and after some rearrangements we easily obtain:

$$\begin{aligned}\Gamma=\left(\frac{N}{2\pi}\right)^{(T+1)r}\int\prod_{\rho t}\mathrm{d}\hat{m}_\rho^t\,\mathrm{d}m_\rho^t\exp\Bigg\{&F(N,\{m_\rho^0\}_\rho)+\mathrm{i}N\sum_{\rho t}\hat{m}_\rho^t m_\rho^t\\&-\sum_{it\neq T}\ln\left[\int\mathrm{d}\sigma\,\mathbf{T}(\sigma,\xi_i,\{m_\rho^t\}_\rho)\right]+\sum_i\ln\left[\int\mathrm{d}\sigma\exp-\mathrm{i}\sum_\rho\hat{m}_\rho^0M^\rho(\sigma,\xi_i)\right]\\&+\sum_{it\neq0}\ln\left[\int\mathrm{d}\sigma\,\mathbf{T}(\sigma,\xi_i,\{m_\rho^{t-1}\}_\rho)\exp-\mathrm{i}\sum_\rho\hat{m}_\rho^tM^\rho(\sigma,\xi_i)\right]+\sum_{\rho t}\lambda_\rho^t m_\rho^t\Bigg\}.\end{aligned} \tag{3.2}$$

Let us define the following function of m_ρ^t and $\hat{m}_\rho^t$ for all values of ρ and t between 0 and T, dependent on a discrete parameter ξ:

$$\begin{aligned}\mathbf{H}(\xi,\{\hat{m}_\rho^t\}_{\rho t},\{m_\rho^t\}_{\rho t})=&\ln\left[\int\mathrm{d}\sigma\exp-\sum_\rho\hat{m}_\rho^0M^\rho(\sigma,\xi)\right]\\&+\sum_{t\neq0}\ln\left[\int\mathrm{d}\sigma\,\mathbf{T}(\sigma,\xi,\{m_\rho^{t-1}\}_\rho)\exp-\sum_\rho\hat{m}_\rho^tM^\rho(\sigma,\xi)\right]\\&-\sum_{t\neq T}\ln\left[\int\mathrm{d}\sigma\,\mathbf{T}(\sigma,\xi,\{m_\rho^t\}_\rho)\right].\end{aligned} \tag{3.3}$$

Introducing this function into (3.2) and performing the change of variables $\hat{m}_\rho^t\to$

$-\mathrm{i}\hat{m}_p^t$, we have for the generating function (from now on, the contour of integration for the variables $\hat{m}_p^t$ is implicitly taken to be $-\mathrm{i}\infty\to+\mathrm{i}\infty$):

$$\Gamma=\left(\frac{-\mathrm{i}N}{2\pi}\right)^{(T+1)r}\int\prod_{pt}\mathrm{d}\hat{m}_p^t\,\mathrm{d}m_p^t\exp\left(\sum_{pt}\lambda_p^t m_p^t\right)$$
$$\exp\left\{F(N,\{m_p^0\}_p)+N\sum_{pt}\hat{m}_p^t m_p^t+\sum_i \mathbf{H}(\xi_i,\{\hat{m}_p^t\}_{pt},\{m_p^t\}_{pt})\right\}. \tag{3.4}$$

The contribution of $\mathbf{H}$ to the integral is in the form of a sum of values of $\mathbf{H}$, considered as a function of $\hat{m}_p^t$ and m_p^t, for each ξ_i. As ξ takes only a finite number of values, the possible values of $\mathbf{H}$ span a finite-dimensional space. Therefore, using the fact that for each i, ξ_i is a random independent variable, we can use the central limit theorem to approximate this sum, up to the first order in $\sqrt{N}$:

$$\sum_i \mathbf{H}(\xi_i,\{\hat{m}_p^t\}_{pt},\{m_p^t\}_{pt})=N\langle\mathbf{H}(\xi,\{\hat{m}_p^t\}_{pt},\{m_p^t\}_{pt})\rangle_{k(\xi)}+\sqrt{N}\mathbf{Z}(\{\hat{m}_p^t\}_{pt},\{m_p^t\}_{pt}) \tag{3.5}$$

where $\mathbf{Z}$ is a random function belonging to the same finite-dimensional space as $\mathbf{H}$, obeying a Gaussian probability with zero mean and covariance given by:

$$\langle\mathbf{Z}(\{\hat{m}_p^t\}_{pt},\{m_p^t\}_{pt})\mathbf{Z}(\{\hat{n}_p^t\}_{pt},\{n_p^t\}_{pt})\rangle=\langle\mathbf{H}(\xi,\{\hat{m}_p^t\}_{pt},\{m_p^t\}_{pt})\mathbf{H}(\xi,\{\hat{n}_p^t\}_{pt},\{n_p^t\}_{pt})\rangle_{k(\xi)}$$

and $\langle\rangle_{k(\xi)}$ denotes the mean value relative to the density of probability k. $\mathbf{Z}$ describes, in fact, the fluctuations up to the first order of $1/\sqrt{N}$ due to the random nature of ξ as opposed to the 'thermal' fluctuations due to the stochastic nature of the evolution (2.2) that we will evaluate later. It must be observed that the two are different in nature, and that if the former can be calculated independently of the latter, the inverse is not *a priori* true.

Finally we have

$$\Gamma=\left(\frac{-\mathrm{i}N}{2\pi}\right)^{(T+1)r}\int\prod_{pt}\mathrm{d}\hat{m}_p^t\,\mathrm{d}m_p^t\exp\left(\sum_{pt}\lambda_p^t m_p^t\right)$$
$$\times\exp\left\{F(N,\{m_p^0\}_p)+N\sum_{pt}\hat{m}_p^t m_p^t+N\langle\mathbf{H}(\xi,\{\hat{m}_p^t\}_{pt},\{m_p^t\}_{pt})\rangle_{k(\xi)}\right.$$
$$\left.+\sqrt{N}\mathbf{Z}(\{\hat{m}_p^t\}_{pt},\{m_p^t\}_{pt})\right\}. \tag{3.6}$$

At this stage we can use a saddle-point method to further approximate the generating function, but a slightly unusual method, as the saddle point itself depends on N. When N tends to infinity, the argument of the second exponential diverges (the first exponential remains fixed). It diverges exponentially, and the maximum value along the path of integration diverges compared with the other values, even if its value and position depends on N. For this to be true, even for the term due to F, we choose from now on F to be of the form $NG(\{m_p^0\}_p)$, with a normalization constant added, so that F denotes a density of probability. Another factor is that T must remain fixed with N. Consequently, the results are valid if we observe the evolution for a time very small compared with N (in fact T must be small compared with $\sqrt{N}$). Now we can deform the original contour of integration such that it passes through the saddle point of this argument and corresponds to the path of steepest descent near it. This contour is also chosen so that the absolute maximum of the argument along it is reached only at this point. The saddle point depends on N, but as N tends to infinity (while p and T remain fixed) only the vicinity of this saddle point contributes to the integral along the path of steepest descent (which depends also on N). Let us denote by ${}^*m^t_\rho$, ${}^*\hat{m}^t_\rho$ this saddle point, and by n^t_ρ, $\hat{n}^t_\rho$ the deviation from it of the parameters of integration along the contour.

The saddle-point equations are derived in appendix 1. One important result is that the role of G is to constrain the values of the parameters for $t=0$. If it is correctly chosen, for example if it is a linear function, it can force the values of ${}^*m^0_\rho$. The second result is that, except at time 0, all the parameters ${}^*\hat{m}^t_\rho$ equal 0. And finally, we obtain an equation giving ${}^*m^{t+1}_\rho$ as a function of ${}^*m^t_\rho$. Consequently, if G is correctly chosen, we have only one saddle point. The evolution of the parameters is given by, up to first order in $1/\sqrt{N}$:

$$ {}^*m^{t+1}_\rho = \left\langle \frac{\int \mathrm{d}\sigma\, M^\rho(\sigma,\xi)\mathbf{T}(\sigma,\xi,\{{}^*m^t_\eta\}_\eta)}{\int \mathrm{d}\sigma\, \mathbf{T}(\sigma,\xi,\{{}^*m^t_\eta\}_\eta)} \right\rangle_{k(\xi)} + \frac{1}{\sqrt{N}}\mathbf{Y}_\rho(\{{}^*m^t_\rho\}_\rho) \tag{3.7} $$

where $\mathbf{Y}_\rho$ is a random function, due to the repartition of the ξ_i, obeying a Gaussian probability with zero mean and correlations (the first mean value is here relative to the probability of repartition of all the ξ_i):

$$ \langle \mathbf{Y}_\rho(\{m_\eta\}_\eta)\mathbf{Y}_{\rho'}(\{m'_\eta\}_\eta)\rangle = \left\langle \frac{\int \mathrm{d}\sigma\, M^\rho(\sigma,\xi)\mathbf{T}(\sigma,\xi,\{m_\eta\}_\eta)}{\int \mathrm{d}\sigma\, \mathbf{T}(\sigma,\xi,\{m_\eta\}_\eta)} \frac{\int \mathrm{d}\sigma\, M^{\rho'}(\sigma,\xi)\mathbf{T}(\sigma,\xi,\{m'_\eta\}_\eta)}{\int \mathrm{d}\sigma\, \mathbf{T}(\sigma,\xi,\{m'_\eta\}_\eta)} \right\rangle_{k(\xi)}. \tag{3.8} $$

We can approximate the generating function, using the Taylor expansion of the argument of the second exponential in (3.6) and dropping irrelevant terms (terms not of order N in the exponential, whose contribution is negligible for the determination of the fluctuations up to the first order in $1/\sqrt{N}$) by:

$$ \begin{aligned} &\Phi \exp\left(\sum_{\rho t} \lambda^t_\rho {}^*m^t_\rho\right) \int_{C_N} \prod_{\rho t} \mathrm{d}n^t_\rho\, \mathrm{d}\hat{n}^t_\rho \exp\left(\sum_{\rho t} \lambda^t_\rho n^t_\rho\right) \\ &\quad \times \exp\left\{ \frac{N}{2}\sum_{\rho\rho'} n^0_\rho n^0_{\rho'} \frac{\partial^2 G}{\partial m_\rho \partial m_{\rho'}}(\{{}^*m^0_\eta\}_\eta) + N\sum_{\rho t} \hat{n}^t_\rho n^t_\rho \right. \\ &\quad + \frac{N}{2}\sum_{\rho\rho' tt'} n^t_\rho n^{t'}_{\rho'} \left\langle \frac{\partial^2 \mathbf{H}}{\partial m^t_\rho \partial m^{t'}_{\rho'}}(\xi,\{{}^*\hat{m}^\tau_\eta\}_{\eta\tau},\{{}^*m^\tau_\eta\}_{\eta\tau}) \right\rangle_{k(\xi)} \\ &\quad + N\sum_{\rho\rho' tt'} \hat{n}^t_\rho n^{t'}_{\rho'} \left\langle \frac{\partial^2 \mathbf{H}}{\partial \hat{m}^t_\rho \partial m^{t'}_{\rho'}}(\xi,\{{}^*\hat{m}^\tau_\eta\}_{\eta\tau},\{{}^*m^\tau_\eta\}_{\eta\tau}) \right\rangle_{k(\xi)} \\ &\quad \left. + \frac{N}{2}\sum_{\rho\rho' tt'} \hat{n}^t_\rho \hat{n}^{t'}_{\rho'} \left\langle \frac{\partial^2 \mathbf{H}}{\partial \hat{m}^t_\rho \partial \hat{m}^{t'}_{\rho'}}(\xi,\{{}^*\hat{m}^\tau_\eta\}_{\eta\tau},\{{}^*m^\tau_\eta\}_{\eta\tau}) \right\rangle_{k(\xi)} \right\} \end{aligned} \tag{3.9} $$

where Φ is dependent on $^*m_\rho^t$, but not on λ_ρ^t. Its value can also be obtained through the observation that for λ_ρ^t null, the generating function must be equal to 1. C_N is the contour of steepest descent near the saddle point, translated so that this point corresponds to 0. As only the vicinity of the saddle point contributes to the generating function when N tends to infinity, this contour can be chosen so that it corresponds to the path of steepest descent of the Gaussian function in the above equation (obviously tangent to the real steepest descent path at 0). In fact it is equivalent to choose any linear contour for which 0 is an absolute maximum, implying that the function decreases exponentially away from 0. By derivation of the generating function it is easy to obtain the interpretation of $^*m_\rho^t$ at the saddle point: it is in fact the 'thermal' average, or average relative to the stochastic nature of the evolution, of the parameters of the network defined by (2.1). Equation (3.7) describes the evolution of these averages, up to the order 1 in $1/\sqrt{N}$. The values of the $^*m_\rho^0$ are then simply the mean value of the parameters for the probability given by G. These are not simply expressed in terms of G because G is a function of the m_ρ^t, and the number of states of the network with given parameters depends highly on these parameters (most of the states have parameters near 0).

By derivation of the generating function and using the remark above concerning the value of Φ, we obtain the correlations of the 'thermal' flucutations δm_ρ^t and $\delta\hat{m}_\rho^t$ of the order parameters of the network, m_ρ^t and $\hat{m}_\rho^t$, near their thermal averages and up to the first order of $1/\sqrt{N}$. The result is for the correlation between the fluctations δm_ρ^t (the other correlations have exactly the same form):

$$\langle \delta m_\rho^t \delta m_{\rho'}^{t'} \rangle = \frac{1}{D} \int_{C_N} \prod_{\eta\tau} \mathrm{d}\hat{n}_\eta^\tau \, \mathrm{d}n_\eta^\tau \, n_\rho^t n_{\rho'}^{t'}$$

$$\times \exp\left\{ \frac{N}{2} \sum_{\eta\eta'} n_\eta^0 n_{\eta'}^0 \frac{\partial^2 G}{\partial m_\eta \partial m_{\eta'}} (\{^*m_\kappa^0\}_\kappa) + N \sum_{\eta\tau} \hat{n}_\eta^\tau n_\eta^\tau \right.$$

$$+ \frac{N}{2} \sum_{\eta\eta'\tau\tau'} n_\eta^\tau n_{\eta'}^{\tau'} \left\langle \frac{\partial^2 \mathbf{H}}{\partial m_\eta^\tau \partial m_{\eta'}^{\tau'}} (\xi, \{^*\hat{m}_\kappa^u\}_{\kappa u}, \{^*m_\kappa^u\}_{\kappa u}) \right\rangle_{k(\xi)}$$

$$+ N \sum_{\eta\eta'\tau\tau'} \hat{n}_\eta^\tau n_{\eta'}^{\tau'} \left\langle \frac{\partial^2 \mathbf{H}}{\partial \hat{m}_\eta^\tau \partial m_{\eta'}^{\tau'}} (\xi, \{^*\hat{m}_\kappa^u\}_{\kappa u}, \{^*m_\kappa^u\}_{\kappa u}) \right\rangle_{k(\xi)}$$

$$\left. + \frac{N}{2} \sum_{\eta\eta'\tau\tau'} \hat{n}_\eta^\tau \hat{n}_{\eta'}^{\tau'} \left\langle \frac{\partial^2 \mathbf{H}}{\partial \hat{m}_\eta^\tau \partial \hat{m}_{\eta'}^{\tau'}} (\xi, \{^*\hat{m}_\kappa^u\}_{\kappa u}, \{^*m_\kappa^u\}_{\kappa u}) \right\rangle_{k(\xi)} \right\} \tag{3.10}$$

where D is the same Gaussian integral but without the term $n_{\rho'}^t n_{\rho'}^{t'}$ outside the exponential (D is a simple noirmalization factor). In this expression the Gaussian function in the two integrals is defined by a real quadratic form in the exponential (the C_N is complex, of course). Introducing the linear change of variable which brings the contour C_N on the real axis and calculating the Gaussian integral, it is easy to show that these correlations can be collectively expressed as the inverse of twice the opposite of the matrix defining the above quadratic form (which is, as the remark above implies, a real symmetrical matrix). This inverse is studied in appendix 2. The elements of this inverse can be obtained recursively, starting from $t=0$ up to $t=T$. The fluctuations for $t=0$ are constrained by G. In turn, the second differential of G at the saddle point must obey a limiting relation, otherwise the correlations of the fluctuations obtained at time 0 correspond to a negative symmetrical matrix, which is absurd. The case of G linear mentioned above does not violate this condition. If we concentrate on the 'thermal' correlations between the different δm_ρ^t for the same time, we have the following relation (see appendix 2):

$$\langle \delta m_\rho^{t+1} \delta m_{\rho'}^{t+1} \rangle = K_{\rho\rho'}(\{^*m_\kappa^t\}_\kappa) + \sum_{\eta\eta'} L_{\rho\eta}(\{^*m_\kappa^t\}_\kappa) \langle \delta m_\eta^t \delta m_{\eta'}^t \rangle L_{\rho'\eta'}(\{^*m_\kappa^t\}_\kappa) \tag{3.11}$$

where $K_{\rho\rho'}$ and $L_{\rho\rho'}$ are functions of the $^*m_\rho^t$ and are defined by

$$K_{\rho\rho'}(\{m_\eta\}_\eta) = \left\langle \frac{\int d\sigma\, M^\rho(\sigma,\xi) M^{\rho'}(\sigma,\xi) \mathbf{T}(\sigma,\xi,\{m_\eta\}_\eta)}{\int d\sigma\, \mathbf{T}(\sigma,\xi,\{m_\eta\}_\eta)} - \frac{\int d\sigma\, M^\rho(\sigma,\xi)\mathbf{T}(\sigma,\xi,\{m_\eta\}_\eta) \int d\sigma\, M^{\rho'}(\sigma,\xi)\mathbf{T}(\sigma,\xi,\{m_\eta\}_\eta)}{\left[\int d\sigma\, \mathbf{T}(\sigma,\xi,\{mk_\eta\}_\eta)\right]^2} \right\rangle_{k(\xi)} \tag{3.12a}$$

$$L_{\rho\rho'}(\{m_\eta\}_\eta) = \left\langle \frac{\int d\sigma\, M^\rho(\sigma,\xi) \frac{\partial \mathbf{T}}{\partial m_{\rho'}}(\sigma,\xi,\{m_\eta\}_\eta)}{\int d\sigma\, \mathbf{T}(\sigma,\xi,\{m_\eta\}_\eta)} - \frac{\int d\sigma\, M^\rho(\sigma,\xi)\mathbf{T}(\sigma,\xi,\{m_\eta\}_\eta) \int d\sigma \frac{\partial \mathbf{T}}{\partial m_{\rho'}}(\sigma,\xi,\{m_\eta\}_\eta)}{\left[\int d\sigma\, \mathbf{T}(\sigma,\xi,\{m_\eta\}_\eta)\right]^2} \right\rangle_{k(\xi)}. \tag{3.12b}$$

The equations (3.7), (3.8), (3.11) and (3.12) characterize the evolution of the macroscopic parameters of the network and their fluctuations. We can present these results in another way. If we denote by $\bar{m}_\rho^t$ the average of the parameters relative to the stochastic evolution and the repartition of the ξ_i at the same time, and dropping the terms not of the first order in $1/\sqrt{N}$, we can pose:

$$m_\rho^t = \bar{m}_\rho^t + \frac{1}{\sqrt{N}} \delta m_\rho^t + \frac{1}{\sqrt{N}} \Delta m_\rho^t \tag{3.13}$$

where the evolution of $\bar{m}_\rho^t$ is given by

$$\bar{m}_\rho^{t+1} = \left\langle \frac{\int d\sigma\, M^\rho(\sigma,\xi)\mathbf{T}(\sigma,\xi,\{\bar{m}_\eta^t\}_\eta)}{\int d\sigma\, \mathbf{T}(\sigma,\xi,\{\bar{m}_\eta^t\}_\eta)} \right\rangle_{k(\xi)} \tag{3.14}$$

and where δm_ρ^t are the 'thermal' fluctuations, obeying (3.11) with $\bar{m}_\rho^t$ in place of $^*m_\rho^t$

(in fact, as long as the difference is of order $1/\sqrt{N}$, the result is the same up to the same order for the fluctuations). Δm_ρ^t are then what we can call the 'quenched' fluctuations, due to the repartition of the ξ_i. From (3.7) it is easy to show that these fluctuations obey the following evolution:

$$\Delta m_\rho^{t+1} = \mathbf{Y}_\rho(\{\bar{m}_\eta^t\}_\eta) + \sum_{\rho'} L_{\rho\rho'}(\{\bar{\mathrm{m}}_\eta^t\}_\eta)\Delta m_{\rho'}^t. \tag{3.15}$$

We find $L_{\rho\rho'}$ again in this equation because it is the first derivative of the iteration defined by (3.14). This evolution is the equivalent of the equation (3.11) for the 'quenched' fluctuations. It consists of a term at time t characteristic of the averages $\bar{m}_\rho^t$ to which is added the transport of the previous fluctuations due to the derivative of the evolution of these averages. The difference is that the first term is independent of the previous fluctuations for the 'thermal' case, which is not true for the 'quenched' case, so that the exact equivalent of (3.11) in this case would include the correlations of this first term with the previous fluctuations. This final result shows also that in fact the two fluctuations, 'thermal' and 'quenched' are independent. All these results are correct as long as the fluctutations remain small. If this is true initially, then this is true at time T provided T is small compared with $\sqrt{N}$.

4. Applications of the general model

These results can be applied to various cases of synchronous neural networks. For example, for the asymmetrical Little model (see equations (2.3), (2.4) and (2.5)) we obtain the following evolution of the order parameters, which are the averages of the correlations with the stored patterns defined by (2.3):

$$\bar{m}_\mu^{t+1} = \left\langle \xi_\mu \tanh\beta\left[\sum_{\nu\nu'} A_{\nu\nu'}\xi_\nu \bar{m}_{\nu'}^t + \sum_\nu \Theta_\nu \xi_\nu\right]\right\rangle_{k(\xi)}. \tag{4.1}$$

The mean value is relative to the probability of all the values of ξ belonging to $\{-1,+1\}^P$. This is the result obtained in [3]. For the fluctuations due to the stored patterns we have

$$\Delta m_\mu^{t+1} = \mathbf{Y}_\mu(\{\bar{m}_\nu^t\}_\nu) + \beta\sum_\nu L_{\mu\nu}(\{\bar{m}_{\mu'}^t\}_{\mu'})\Delta m_\nu^t \tag{4.2}$$

with

$$L_{\mu\nu}(\{m_{\mu'}\}_{\mu'}) = \left\langle \xi_\mu\left(\sum_{\nu'} A_{\nu\nu'}\xi_{\nu'}\right)\left\{\tanh^2\beta\left[\sum_{\mu'\nu'} A_{\mu'\nu'}\xi_{\mu'}m_{\nu'} + \sum_{\mu'}\xi_{\mu'}\Theta_{\mu'}\right] - 1\right\}\right\rangle_{k(\xi)} \tag{4.3}$$

and where $\mathbf{Y}_\mu$ are random functions with Gaussian law and correlations given by

$$\langle \mathbf{Y}_\mu(\{m_{\mu'}\}_{\mu'})\mathbf{Y}_\nu(\{n_{\mu'}\}_{\mu'})\rangle = \langle \xi_\mu \xi_\nu \tanh\beta\left[\sum_{\mu'\nu'} A_{\mu'\nu'}\xi_{\mu'}m_{\nu'} + \sum_{\mu'} \xi_{\mu'}\Theta_{\mu'}\right] \times \tanh\beta\left[\sum_{\mu'\nu'} A_{\mu'\nu'}\xi_{\mu'}n_{\nu'} + \sum_{\mu'} \xi_{\mu'}\Theta_{\mu'}\right]\rangle_{k(\xi)}. \tag{4.4}$$

The thermal fluctuations obey, in turn, the following evolution equation:

$$\langle \delta m_\mu^{t+1}\delta m_\nu^{t+1}\rangle = \beta^2 \sum_{\mu'\nu'} L_{\mu\mu'}(\{\bar{m}^t_{\mu''}\}_{\mu''})L_{\nu\nu'}(\{\bar{m}^t_{\mu''}\}_{\mu''})\langle \delta m^t_{\mu'}\delta m^t_{\nu'}\rangle + \langle \xi_\mu\xi_\nu\left\{1-\tanh^2\beta\left[\sum_{\mu'\nu'} A_{\mu'\nu'}\xi_{\mu'}\bar{m}^t_{\nu'} + \sum_{\mu'}\xi_{\mu'}\Theta_{\mu'}\right]\right\}\rangle_{k(\xi)}. \tag{4.5}$$

These results can be extended to the nonlinear Little model where the connection matrix is given by

$$J_{ij} = \frac{1}{N} Q(\xi_i, \xi_j) \tag{4.6}$$

where the ξ_i are the vectors belonging to $\{-1, +1\}^p$, the same as those for the normal Little model, and Q is a function of two vectors of this kind. We suppose that the thresholds are null. The global parameters of the evolution are indexed by this type of vector and are defined by

$$m_\xi = \frac{1}{N}\sum_i \sigma_i \prod_\mu \delta_{\xi_\mu \xi^i_\mu}. \tag{4.7}$$

The order parameters are then the averages of these parameters (relative to the stochastic evolution and the random nature of the stored patterns) and we have for their evolution:

$$\bar{m}^{t+1}_\xi = k(\xi)\tanh\beta\left[\sum_{\xi'} Q(\xi, \xi')\bar{m}^t_{\xi'}\right]. \tag{4.8}$$

For the 'quenched' fluctuations we have

$$\Delta m^{t+1}_\xi = \mathbf{Y}_\xi(\{\bar{m}^t_{\xi'}\}_{\xi'}) + \beta\sum_{\xi'} L_{\xi\xi'}(\{\bar{m}^t_{\xi''}\}_{\xi''})\Delta m^t_{\xi'} \tag{4.9}$$

with the following correlations for the random functions with Gaussian law $\mathbf{Y}_\xi$:

$$\langle \mathbf{Y}_\xi(\{m_{\xi''}\}_{\xi''})\mathbf{Y}_{\xi'}(\{n_{\xi''}\}_{\xi''})\rangle = k(\xi)\delta_{\xi\xi'}\tanh\beta\left[\sum_{\xi''} Q(\xi, \xi'')m_{\xi''}\right]\tanh\beta\left[\sum_{\xi''} Q(\xi, \xi'')n_{\xi''}\right] \tag{4.10}$$

and for $L_{\xi\xi'}$

$$L_{\xi\xi'}(\{m_{\xi''}\}_{\xi''}) = k(\xi)Q(\xi, \xi')\left\{\tanh^2\beta\left[\sum_{\xi''} Q(\xi, \xi'')m_{\xi''}\right] - 1\right\}. \tag{4.11}$$

For the 'thermal' fluctuations we have

$$\langle \delta m_{\xi}^{t+1} \delta m_{\xi'}^{t+1} \rangle = k(\xi)\delta_{\xi\xi'}\left\{1-\tanh^2\beta\left[\sum_{\xi''} Q(\xi,\xi'')\bar{m}_{\xi''}^{t}\right]\right\}$$

$$+\beta^2 \sum_{\chi\chi'} L_{\xi\chi}(\{\bar{m}_{\xi''}^{t}\}_{\xi''}) L_{\xi'\chi'}(\{\bar{m}_{\xi''}^{t}\}_{\xi''}) \langle \delta m_{\chi}^{t} \delta m_{\xi'}^{t} \rangle. \tag{4.12}$$

Another possible generalization of the Little model is to have, instead of spin-like neurons with two value, Potts spins for the states σ_i. In this case, there are q states that we will denote by γ_k corresponding to q vectors of $\mathbb{R}^{q-1}$ with scalar product

$$\gamma_k \cdot \gamma_{k'} = q\delta_{kk'} - 1. \tag{4.13}$$

The set of these vectors will be denoted by Ω. The connections are then determined by p stored patterns ξ_i^{μ} belonging to Ω by

$$J_{ij} = \frac{1}{N}\sum_{\mu\nu} A_{\mu\nu}\xi_i^{\mu} \cdot \xi_j^{\nu} \tag{4.14}$$

and the transition function by

$$\mathbf{T} = \exp\beta\left[\sum_j J_{ij}\boldsymbol{\sigma}_i^{t+1} \cdot \boldsymbol{\sigma}_j^{t}\right] \tag{4.15}$$

the parameters of the evolution are then

$$n_{\mu}(x,y) = \frac{1}{N}\sum_i (x \cdot \xi_i^{\mu})(y \cdot \sigma_i) \tag{4.16}$$

with $xy\sigma$ and ξ belonging to Ω. We have consequently the following evolution for their averages.

$$n_{\mu}^{t+1}(x,y) = \left\langle \frac{\sum_{\sigma} (x \cdot \xi_{\mu})(y \cdot \sigma) \exp\beta\left[\sum_{\nu\nu'} A_{\nu\nu'} n_{\nu'}^{t}(\xi_{\nu}, \sigma)\right]}{\sum_{\sigma} \exp\beta\left[\sum_{\nu\nu'} A_{\nu\nu'} n_{\nu'}^{t}(\xi_{\nu}, \sigma)\right]} \right\rangle_{k(\xi)} \tag{4.17}$$

with k the probability on Ω^p, ξ an element of Ω^p with components ξ_{μ}. We do not pose the equations for the fluctuations as they are very cumbersome. But these can be obtained easily from the general formulas (3.11), (3.12), (3.14), (3.15) and (3.8).

5. Conclusions

The main results of this article are the equations (3.14), (3.11), (3.12), (3.15) and (3.8). They describe the evolution of the order parameters and their fluctuations. The important point is that there are two kinds of fluctuations, independent of each other: fluctutations due to the stochastic nature of the evolution, and the 'quenched' fluctuations, due to the random nature of the choice of the ξ, which correspond in most neural networks models to the stored patterns. A certain number of essential ingredients are needed for these results to be valid. At the present stage, it is restricted to synchronous models. To extend the method to asynchronous neural networks seems to be possible but would require the use of path integrals. The hypothesis of identical probability for each ξ_i is not, in fact, necessary for the central limit theorem. It can be proven given a certain number of conditions on the moments of the probabilities of each ξ_i [7]. Even the hypothesis of independence of the probabilities for each i is not an absolute necessity. Concepts of conditioning can be used to extend the validity of the results to dependent cases (see [7]). Nonetheless, a fundamental restriction is that the set of possible values of ξ_i is finite when N tends to infinity, which precludes the study of the case of the number of stored patterns growing with N, the number of neurons. In conclusion, the method we use to solve the general model of synchronous neural networks presented in section 2 seems to be extendable to many other cases as well. Its main limitation is that it does not seem to be applicable to the study of the saturation of neural networks.

Appendix 1

Let us pose the following definitions:

$$a_t = \ln\left[\int \mathrm{d}\sigma\,\mathbf{T}(\sigma, \xi, \{m_\rho^t\}_\rho)\right]$$

$$b_t = \ln\left[\int \mathrm{d}\sigma\,\mathbf{T}(\sigma, \xi, \{m_\rho^t\}_\rho)\exp - \sum_\rho \hat{m}_\rho^{t+1} M^\rho(\sigma, \xi)\right]$$

$$c = \ln\left[\int \mathrm{d}\sigma \exp - \sum_\rho \hat{m}_\rho^0 M^\rho(\sigma, \xi)\right]. \tag{A1.1}$$

H is then defined as:

$$\mathbf{H} = c + \sum_{t\neq T} b_t + \sum_{t\neq T} a_t \tag{A1.2}$$

where a_t depend on the m_ρ^t only, b_t on the m_ρ^t and the $\hat{m}_\rho^{t+1}$ only, and c on the $\hat{m}_\rho^0$. As **Z** belong to the same space as **H**, it has the same structure, and we have

$$\mathbf{Z} = c' + \sum_{t\neq T} b_t' + \sum_{t\neq T} a_t' \tag{A1.3}$$

with the same dependencies. The saddle point equations are then (divided by N): by derivation for m_ρ^0, for all possible values of ρ

$$\frac{\partial g}{\partial m_\rho}(\{{}^*m_{\rho'}^0\}_{\rho'}) + {}^*\hat{m}_\rho^0 - \left\langle \frac{\partial a_0}{\partial m_\rho^0}\right\rangle_{k(\xi)} + \left\langle \frac{\partial b_0}{\partial m_\rho^0}\right\rangle_{k(\xi)} - \frac{1}{\sqrt{N}}\frac{\partial a_0'}{\partial m_\rho^0} + \frac{1}{\sqrt{N}}\frac{\partial b_0'}{\partial m_\rho^0} = 0 \tag{A1.4}$$

by derivation for m_ρ^T

$${}^*\hat{m}_\rho^T = 0 \tag{A1.5}$$

and for $m_\rho^t (t \in]0, T[\cap \mathbb{N})$

$$ {}^*\hat{m}_\rho^t - \left\langle \frac{\partial a_t}{\partial m_\rho^t} \right\rangle_{k(\xi)} + \left\langle \frac{\partial b_t}{\partial m_\rho^t} \right\rangle_{k(\xi)} - \frac{1}{\sqrt{N}} \frac{\partial a_t'}{\partial m_\rho^t} + \frac{1}{\sqrt{N}} \frac{\partial b'}{\partial m_\rho^t} = 0 \tag{A1.6} $$

by derivation for $\hat{m}_\rho^0$

$$ {}^*m_\rho^0 = -\left\langle \frac{\partial c}{\partial \hat{m}_\rho^0} \right\rangle_{k(\xi)} = \frac{1}{\sqrt{N}} \frac{\partial c'}{\partial \hat{m}_\rho^0} \tag{A1.7} $$

and for $\hat{m}_\rho^t$, $t \neq 0$

$$ {}^*m_\rho^t = -\left\langle \frac{\partial b_{t-1}}{\partial \hat{m}_\rho^t} \right\rangle_{k(\xi)} - \frac{1}{\sqrt{N}} \frac{\partial b_{t-1}'}{\partial \hat{m}_\rho^t}. \tag{A1.8} $$

One important relation is that a_t equals b_t if ${}^*\hat{m}_\rho^{t+1}$ is null. This relation is therefore also true of a_t' and b_t'. From (A1.6) we see that that in that case ${}^*\hat{m}_\rho^t$ is also null, provided that t belongs to $]0, T[\cap \mathbb{N}$. Then from (A1.5) we conclude that, for t different from 0, ${}^*m_\rho^t$ is null. (A1.4) and (A1.7) together give the relations obeyed by the parameters at time 0. If g is linear, for example, (A1.4) gives directly the value of ${}^*\hat{m}_\rho^0$, which in turn gives the value of ${}^*m_\rho^0$ through (A1.7). The final result, the evolution equation for the parameters, is then given by (A1.8). This is in fact identical to (3.7) where $\mathbf{Y}$ is nothing but the derivative of b_t relative to ${}^*m_\rho^t$ for $\hat{m}_\rho^t$ null. (3.8) is then given by the correlations of b' obtained from the correlations of $\mathbf{Z}$, in turn given by the central limit theorem in terms of the correlations of $\mathbf{H}$.

Appendix 2

To calculate the inverse of the matrix, let us pose the following partial matrices (the index ρ is implicit in these definitions):

$$ G = \frac{\partial^2 g}{(\partial m^t)^2} $$

$$ \Gamma^t = \left\langle \frac{\partial^2 \mathbf{H}}{(\partial \hat{m}^t)^2} \right\rangle_{k(\xi)} $$

$$ H^t = \left\langle \frac{\partial^2 \mathbf{H}}{\partial \hat{m}^{t+1} \partial m^t} \right\rangle_{k(\xi)}. \tag{A2.1} $$

We also have the following relation at the saddle point, due to the structure of $\mathbf{H}$:

$$ \left\langle \frac{\partial^2 \mathbf{H}}{\partial \hat{m}^t \partial m^t} \right\rangle_{k(\xi)} = I \text{ (identity)}. \tag{A2.2} $$

These matrices correspond, in fact, to elements of the opposite of the matrix to inverse. The other elements of this matrix are null. Let us use the following notation for the wanted correlations, collectively expressed as a matrix of the same type as the one defining the quadratic form in (3.10):

$A_{tt'}$ for the term corresponding to $m^t m^{t'}$
$B_{tt'}$ for the term corresponding to $\hat{m}^t m^{t'}$
$C_{tt'}$ for the term corresponding to $m^t \hat{m}^{t'}$
$D_{tt'}$ for the term corresponding to $\hat{m}^t \hat{m}^{t'}$.

We know that this matrix defining the correlations of the 'thermal' fluctuations must be symmetrical which implies

$$A_{tt'} = {}^T A_{t't} \qquad D_{tt'} = {}^T D_{t't} \qquad B_{tt'} = {}^T C_{t't}. \tag{A2.3}$$

If we multiply this matrix to the right of the matrix to inverse we must obtain the identity. As the two matrices are symmetrical, this implies automatically the same result for the product to the left. If we find a matrix obeying this relation, this is then the wanted inverse, and the solution is certainly unique. This relation translates in a number of relations for the partial matrices defined above.

From these relations, it can be shown that the inverse can indeed be calculated, under certain conditions for G. If these conditions are not verified, the inverse does not exist. In fact, if a more general condition is not verified the fluctuations at time 0 are not defined (they are not of the first order in $1/\sqrt{N}$). To illustrate how these results are obtained, we derive now the equation (3.11) and the condition for the existence of the inverse from the relations for the partial matrices (the complete derivation of the inverse goes along the same lines but is quite cumbersome, so we will not describe it). First, we have for D:

$$D_{tT} = 0 \qquad \forall t \qquad \text{and} \qquad {}^T H_{t+1} D_{t+1\,t'} + D_{tt'} = 0 \qquad \forall t \neq 0, T$$

which immediately gives

$$\forall t > 0 \qquad D_{tt'} = 0. \tag{A2.4}$$

For B we have

$$B_{Tt'} = -\delta_{t'T} I \qquad \forall t'$$
$${}^T H_{t+1} B_{t+1\,t'} + B_{tt'} = -\delta_{tt'} I \qquad \forall t \neq 0$$

and then by recursion and using the relation between B and C in (A2.3) we have:

$$\begin{array}{lllll} \forall t' > t & B_{t't} = 0 & \text{and} & B_{tt} = -I & \text{all this for } t \neq 0 \\ \forall t' > T & C_{tt'} = 0 & \text{and} & C_{tt} = -I & \text{all this for } t \neq 0. \end{array} \tag{A2.5}$$

Using this we have the following equations for A:

$$A_{tt} - \Gamma_t + H_t A_{t-1\,t} = 0 \qquad \forall t$$
$$A_{tt'} + H_t A_{t-1\,t'} = 0 \qquad \forall t' < T \tag{A2.6}$$

and therefore

$$\forall t \neq 0, T \qquad A_{t+1\,t+1} = \Gamma_{t+1} - H_{t+1} A_{t+1\,t}$$
$$= \Gamma_{t+1} - H_{t+1} {}^T A_{t\,t+1}$$

then we have under the same condition

$$A_{t+1\,t} + H_{t+1} A_{tt} = 0$$
$$\Rightarrow {}^T A_{t+1\,t} = -A_{tt} {}^T H_{t+1}$$

and finally we obtain:

$$A_{t+1\,t+1}=\Gamma_{t+1}+H_{t+1}A_{tt}{}^{T}H_{t+1} \qquad \forall t\neq 0, T. \tag{A2.7}$$

Equation (A2.7) is in fact exactly the equation (3.11). It is important to note that if A_{tt} is a symmetrical positive matrix, then this is also true of $A_{t+1\,t+1}$, as this is true for Γ_t. This is necesary for A_{tt} to make sense as a correlation matrix for the fluctuations. Now if we analyse the case of $t=0$ (initial fluctuations) we have:

$$\begin{aligned}
&{}^{T}H_1A_{10}+GA_{00}+B_{00}=-I\\
&{}^{T}H_1D_{10}+GC_{00}+D_{00}=0\\
&A_{00}+\Gamma_0B_{00}=0\\
&C_{00}+\Gamma_0D_{00}=-I\\
&A_{10}+H_1A_{00}=0.
\end{aligned} \tag{A2.8}$$

Using this we have

$$\begin{aligned}
&A_{00}=-\Gamma_0B_{00}\\
&C_{00}=-\Gamma_0D_{00}-I\\
&A_{10}=-H_1A_{00}\\
&\Rightarrow {}^{T}H_1H_1\Gamma_0B_{00}-G\Gamma_0B_{00}+B_{00}=-I\\
&\Rightarrow A_{00}=-(\Gamma_0^{-1}+{}^{T}H_1H_1-G)^{-1}.
\end{aligned} \tag{A2.8}$$

The existence of A_{00}, which implies that the right term is inversible, is the necessary condition for the total matrix to have an inverse. For this inverse to make sense, it is also necessary to impose A_{00} symmetrical positive (this will impose the same conditions for other times due to (A2.7)) which gives the necessary condition for C. If it is not verified, then the obtained correlations of the fluctuations at time 0 are negative, which is absurd. In fact, in this case, they are not of order $1/\sqrt{N}$. In the case of g linear, G is null and the condition is verified.

References

[1] Little W A 1974 *Math. Biosci.* **19** 101
[2] Little W A and Shaw G L 1978 *Math. Biosci.* **39** 281
[3] Bernier O 1991 *Europhys. Lett.* **16** 531
[4] Bolle D and Mallezie F 1989 *J. Phys. A: Math. Gen.* **22** 4409
[5] Patrick A E, Picco P, Ruiz J and Zagrebnov V A 1991 *J. Phys. A: Math. Gen.* **24** L637
[6] Shiino M 1990 *J. Stat. Phys.* **59** 1051
[7] Loeve M 1963 *Probability Theory, The University Series in higher mathematics* (Amsterdam: Van Nostrand)