INVESTIGACIÓN REVISTA MEXICANA DE FÍSICA **48** (4) 310–316 AGOSTO 2002

# Stationary processes and equilibrium states in non-symmetric neural networks

A. Castellanos
*Departamento de Física, Universidad de Sonora*
*Apdo. Post. 1626, Hermosillo 83000, Son. México.*
*acastell@fisica.uson.mx*

L. Viana
*Depto. de Física Teórica, CCMC - UNAM*
*Apdo. Post. 2681, 22800 Ensenada B.C., México.*
*laura@ccmc.unam.mx*

Recibido el 9 de noviembre de 2001; aceptado el 4 de abril de 2002

Stationary processes and equilibrium states are discussed in finite separable recurrent neural networks with sequential dynamics and away from saturation. We describe thermal fluctuations of the dynamical order parameters originated as finite size effects of order $\mathcal{O}\ \ N^{-1/2}$ by means of their corresponding Fokker-Planck equation, and find their time dependent probability distribution. We introduce the concept of extended entropy of fluctuations in order to find a general condition to characterize stationary states in Neural Networks with non symmetric interactions. Divergence and rotational of the probability current in the space of fluctuations are also used to differentiate between stationary and equilibrium states. Besides, algebraic conditions are found to know when stationary states can exist. The results are illustrated by analyzing a neural network with a macroscopic dynamical fixed point but not satisfying detailed balance at microscopic level.

*Keywords:* Statistical physics; thermodynamics and nonlinear dynamical systems; neural networks; fuzzy logic; artificial intelligence; stochastic processes; probability theory; stochastic processes and statistics.

Se discuten los procesos estacionarios y los estados de equilibrio en redes neuronales recurrentes, finitas y separables, con dinámica secuencial y muy lejos de la saturación. Por medio de la correspondiente ecuación de Fokker-Planck, describimos las fluctuaciones térmicas de la dinámica de los parámetros de orden, originadas como efectos de tamaño finito de orden ($N^{-1/2}$) y encontramos la distribución de probabilidad dependiente del tiempo. Introducimos el concepto de entropía extendida de las fluctuaciones para encontrar una condición general que caracterice los estados estacionarios en Redes Neuronales con interacciones no simétricas. También se utilizan la divergencia y el rotacional de la corriente de probabilidad en el espacio de fluctuaciones para diferenciar entre estados estacionarios y de equilibrio. Además, se encuentran las condiciones algebraicas para saber cuando pueden existir los estados estacionarios. Los resultados son ilustrados mediante el análisis de una red neuronal con un punto fijo en la dinámica macroscópica pero que no satisface el balance detallado a nivel microscópico.

*Descriptores:* Física estadística; termodinámica y sistemas dinámicos no lineales; redes neuronales; lógica difusa; inteligencia artificial; procesos estocásticos; teoría de la probabilidad; procesos estocásticos y estadística.

PACS: 05.; 07.05.Mh; 02.50.Ey; 02.50.-r

## 1. Introduction

Early works in statistical physics of Neural Networks (NN) were concentrated in the mean field study of equilibrium properties of infinite systems with symmetric interactions (for a review see Ref. 1). However, one of the basic characteristics of NN is their high number of barriers and valleys in the 'energy space', which makes the study of the dynamics essential. In the case of NN with non symmetric interactions, dynamical techniques are in fact the only tool available, as in this kind of systems detailed balance does not hold and the usual statistical concepts such as free energy cannot even be defined.

On the other hand, models of the mean field type, have predictions valid for infinitely large systems ($N \to \infty$). Taking the thermodynamical limit has proved to be a useful approximation to study gases, liquids and solids, where the number of elements is of the order of $10^{23}$; however, the inclusion of finite size effects is clearly important when considering finite systems like the brain having $10^{10}$ neurons working in small groups whose size ranges from a few thousands to $10^5$ neurons. These finite size effects have been shown to play an important role in both symmetric and non symmetric NN [2, 3], even affecting, in some particular cases, the macroscopic behaviour of some systems [4].

Although considering finite size effects and non symmetric interactions is not enough to obtain a good description of biological NN, it is clearly a first and important step in this direction; other characteristics should be included into the models (see Ref. 5 for a good review) such as spatial structure [6–9, 5], as well as short and long-range connectivity [10, 11]. However these considerations are out of the scope of this paper.

Until now, most work in asymmetric neural networks has been concentrated in the study of sequential retrieval of embedded patterns [12], and in the storage capacity of those networks [13], for infinite systems. We believe the study of the stochastic behaviour resulting as a finite size effect in both symmetric and non symmetric NN will contribute towards a better understanding of different aspects of the dynamics of NN and therefore it deserves special attention.

The objective of this work is to study the difference between stationary and equilibrium states from the analysis of the fluctuations originated as finite size effects. Our formalism describes the dynamics of NN to first non trivial order in the system size [15, 16], by separating systematically the $N = \infty$ from the finite N contributions in the Kramers-Moyal expansion [17] and keeping the first non trivial order in $(1/N)$. In this way it is obtained a Fokker-Planck Equation (FPE) which can be written in terms of the fluctuations of the dynamical order parameters, from whose solution it can be constructed the probability distribution for these fluctuations. This formalism does not require symmetry of the interaction matrix, as opposed to other methods [14]; so dynamical flow diagrams for either symmetric or non-symmetric interactions can be obtained by solving numerically the deterministic mean field equations for the fluctuation moments. In the particular case of symmetric interactions the approach to equilibrium is warranted by knowing Lyapunov functions.

We know that any system in thermodynamical equilibrium must obey detailed balance. However, NN with non symmetric interactions can never obey detailed balance; this is reflected on the NN behaviour as the lack of compensation between positive and negative random fluctuations in the order parameters, in a way that induces a rotation of the probability current in the space of the fluctuation variables. In this sense, the divergence and rotational of the probability current are useful tools to characterize the stationary and equilibrium states. In addition to this geometric picture, we obtain the algebraic conditions under which stationary states can exist. We also introduce the concept of 'extended entropy' *à la* Boltzmann, for asymmetric NN, to characterize the stationary states. The limiting (maximum) value of this function allows us to obtain a condition that has to be satisfied by a NN in a stationary state; we find out that this condition is met, in particular, by any symmetric NN in equilibrium, so we can see the equilibrium state as a particular case of a stationary process. These results are illustrated by considering a peculiar system previously analyzed: a NN with a macroscopic fixed point but unable to satisfy detailed balance at the microscopic level [16].

The paper is organized as follows: in Sec. 2, it is presented an overview of the general formalism (for details see Refs. 15, 16); in Sec. 3, it is introduced the concept of extended entropy in order to discuss statistical equilibrium and stationary processes; in Sec. 4, the lack of detailed balance and its relation to the rotation of the probability current are examined to obtain a geometric picture of this phenomenon and the algebraic conditions for the existence of stationary states; finally, in Sec. 5, it is discussed a NN with a fixed point but without detailed balance as an example to illustrate the previous sections.

Our work can be applied to study finite size effects in any kind of finite, separable recurrent NN with sequential dynamics and away from saturation, provided that transference functions are sigmoidal. This includes models with either symmetric or asymmetric interactions, and even models where synapses are state dependent [18, 19]. The analysis of NN with synchronous updating [19, 20] is outside the limits of this theory.

## 2. Finite size effects. General formalism

Let us consider a system composed by a large, but finite, number $N$ of interconnected neurons $\boldsymbol{\sigma} \in \{-1, 1\}^N$, so the vector $\boldsymbol{\sigma}(t) = (\sigma_1(t), \sigma_2(t), \ldots, \sigma_N(t))$ describes the state of the system at a given time. Each variable $\sigma_i(t)$ evolves in time by describing sequential stochastic alignment of the spins under the action of an external field given by

$$h_i(\boldsymbol{\sigma}) = \sum_j J_{ij}\sigma_j + \theta_i,$$

with

$$J_{ij} = \frac{1}{N} \sum_{\mu\nu=1}^{p} \xi_i^\mu A_{\mu\nu}\, \xi_j^\nu \,(1 - \delta_{ij}),$$

$J_{ij}$ is the (separable) synaptic strength of the connection going from neuron $j$ to neuron $i$, $\theta_i$ is a response threshold, and the auto interactions $J_{ii}$ are excluded. These interactions store a finite number $p \ll N$ of quenched binary patterns $\{\xi_i^\mu\}$ chosen at random with $\mu = 1, \ldots, p$ and the notation $\boldsymbol{\xi}_i = (\xi_i^1, \ldots, \xi_i^p) \in \{-1, 1\}^p$, or alternatively, $\boldsymbol{\xi}^\mu = (\xi_1^\mu, \ldots, \xi_N^\mu) \in \{-1, 1\}^N$. And, as we can see, interactions do not need to be symmetrical. The micro-dynamics of the system is defined by a master equation for the probability density $p_t(\boldsymbol{\sigma})$, given by

$$\frac{d}{dt} p_t(\boldsymbol{\sigma}) = \sum_i \left\{ w_i(F_i\boldsymbol{\sigma}) p_t(F_i\boldsymbol{\sigma}) - w_i(\boldsymbol{\sigma}) p_t(\boldsymbol{\sigma}) \right\}, \qquad (1)$$

with

$$w_i(\boldsymbol{\sigma}) = \frac{1}{2}[1 - \sigma_i \tanh(\beta h_i(\boldsymbol{\sigma}))],$$

where $w_i(\boldsymbol{\sigma})$ is the temperature ($T = \beta^{-1}$) dependent transition probability $\sigma_i(t) \to -\sigma_i(t)$, and $F_i$ is an operator that flips the $i-$th spin. If we define the probability density for the pattern overlaps $\boldsymbol{m}$ as $P_t(\boldsymbol{m}) \equiv \sum_\sigma p_t(\boldsymbol{\sigma})\, \delta\left[\boldsymbol{m} - \boldsymbol{m}(\boldsymbol{\sigma})\right]$ with the usual definition for the overlaps

$$\boldsymbol{m}(\boldsymbol{\sigma}) = (m_1(\boldsymbol{\sigma}), \ldots, m_p(\boldsymbol{\sigma})), \qquad m_\mu = \frac{1}{N} \sum_i^N \xi_i^\mu \sigma_i$$

the master Eq. (1) can be rewritten in terms of the macroscopic dynamical variables $\boldsymbol{m}(\boldsymbol{\sigma})$ by inserting $\delta[\boldsymbol{m} - \boldsymbol{m}(\boldsymbol{\sigma})]$ into Eq. (1), and making a Taylor expansion in powers of the vector $2\sigma_i\boldsymbol{\xi}_i/N$. If we expand the resulting Kramers-Moyal equation in powers of $(1/N)$ and keep the two leading orders [23], we obtain a FPE valid on finite time-scales (*i.e.* not scaling with $N$) for the stochastic vector $\boldsymbol{m}(\boldsymbol{\sigma})$. We can now write $\boldsymbol{m}(\boldsymbol{\sigma})$ as the sum of a deterministic part $\boldsymbol{m}^*(t)$ which corresponds to the infinite system, and the leading order stochastic contribution $\boldsymbol{q}(t)/\sqrt{N}$ vanishing as $N \to \infty$

$$\boldsymbol{m} = \boldsymbol{m}^*(t) + \frac{1}{\sqrt{N}} \boldsymbol{q}(t) + \ldots$$

we can now separate the dynamical equations for these two order parameters. In this way, $\boldsymbol{m}^*(t)$ is the deterministic solution to the Liouville equation given by

$$\frac{d}{dt}\boldsymbol{m}^*(t) = \langle \boldsymbol{\xi} \tanh\beta\,[\boldsymbol{\xi}\cdot\boldsymbol{A}\boldsymbol{m}^*(t)+\theta]\rangle_{\boldsymbol{\xi},\theta} - \boldsymbol{m}^*(t), \quad (2)$$

where we defined $\langle\cdots\rangle_{\boldsymbol{\xi},\theta}$ as an average over an ensemble of NN characterized by different realizations of patterns. We can now obtain a FPE for the remaining stochastic part $\boldsymbol{q}(\sigma)$ resulting from the finite size effects as follows:

$$\frac{d}{dt}\mathcal{P}_t(\boldsymbol{q}) = \sum_\mu \frac{\partial}{\partial q_\mu}\left\{\mathcal{P}_t(\boldsymbol{q})\,F_\mu[\boldsymbol{q};t]\right\} + \sum_{\mu\nu}\frac{\partial^2}{\partial q_\mu \partial q_\nu}\left\{\mathcal{P}_t(\boldsymbol{q})\,D_{\mu\nu}[\boldsymbol{q};t]\right\}. \quad (3)$$

In this equation the flow term $F_\mu[\boldsymbol{q};t]$ is given by

$$F_\mu[\boldsymbol{q};t] = K_\mu[\boldsymbol{m}^*(t)] + \sum_\nu L_{\mu\nu}[\boldsymbol{m}^*(t)]q_\nu(t), \quad (4)$$

with

$$K_\mu[\boldsymbol{x}] = \lim_{N\to\infty}\sqrt{N}\left\{\langle \xi_\mu \tanh\beta[\boldsymbol{\xi}\cdot\boldsymbol{A}\boldsymbol{x}+\theta]\rangle_{\boldsymbol{\xi},\theta} - \frac{1}{N}\sum_i \xi_i^\mu \tanh\beta\,[\boldsymbol{\xi}_i\cdot\boldsymbol{A}\boldsymbol{x}+\theta_i]\right\}, \quad (5)$$

$$L_{\mu\nu}[\boldsymbol{x}] = \delta_{\mu\nu} - \beta\sum_\lambda \langle \xi_\mu\xi_\lambda \times \left[1-\tanh^2\beta[\boldsymbol{\xi}\cdot\boldsymbol{A}\boldsymbol{x}+\theta]\right]\rangle_{\boldsymbol{\xi},\theta} A_{\lambda\nu}. \quad (6)$$

It is important to notice that $K_\mu(t)$ makes a 'frozen' correction to the flux, that persists even at $T=0$ and depends on the specific pattern configuration $\{\boldsymbol{\xi}^\mu\}$. On the other hand, the diffusion term $D_{\mu\nu}[t]$ of the FPE, Eq. (3), is found to be symmetric, and given by

$$D_{\mu\nu}[t] = \delta_{\mu\nu} - e^{-t}\lim_{N\to\infty}\frac{1}{N}\sum_i \xi_i^\mu\xi_i^\nu\,\sigma_i(0)\,\tanh\beta[\boldsymbol{\xi}_i\cdot\boldsymbol{A}\boldsymbol{m}(t)+\theta_i] - \int_0^t ds\, e^{s-t}\,\langle \xi_\mu\xi_\nu \tanh\beta[\boldsymbol{\xi}\cdot\boldsymbol{A}\boldsymbol{m}^*(s)+\theta]\,\tanh\beta[\boldsymbol{\xi}\cdot\boldsymbol{A}\boldsymbol{m}^*(t)+\theta]\rangle_{\boldsymbol{\xi},\theta}. \quad (7)$$

It is important to notice that the flux is linear in $\boldsymbol{q}(t)$ and the diffusion term is independent from $\boldsymbol{q}(t)$, therefore the process can be identified as a time dependent Ornstein-Uhlenbeck process, whose formal solution is Gaussian:

$$\mathcal{P}_t(\boldsymbol{q}) = \frac{\exp\left\{-\frac{1}{2}\left[\boldsymbol{q}-\langle\boldsymbol{q}\rangle_t\right]\cdot\boldsymbol{\Xi}^{-1}(t)\left[\boldsymbol{q}-\langle\boldsymbol{q}\rangle_t\right]\right\}}{(2\pi)^{\frac{p}{2}}\sqrt{\det\boldsymbol{\Xi}(t)}} \quad (8)$$

with $\boldsymbol{\Xi}$, the time dependent correlation matrix for the fluctuations $\Xi_{\mu\nu} = \langle q_\mu q_\nu\rangle_t - \langle q_\mu\rangle_t\langle q_\nu\rangle_t$, where $<\cdots>_t$ is the average at time $t$ over different evolutions, for a given pattern configuration $\{\xi_i^\mu\}$. Therefore, the process can be fully characterized by its two first statistical moments, which are found to evolve in time according to the following deterministic equations

$$\frac{d}{dt}\langle\boldsymbol{q}\rangle_t + \langle F_\mu[\boldsymbol{q};t]\rangle_t = 0 \quad (9)$$

$$\frac{d}{dt}\boldsymbol{\Xi}(t) = -\boldsymbol{L}(t)\boldsymbol{\Xi}(t) - \boldsymbol{\Xi}(t)\boldsymbol{L}^T(t) + 2\boldsymbol{D}(t). \quad (10)$$

In order to arrive to explicit expressions for the flux in Eq. (4) and difussion in Eq. (7) terms, for any particular system considered, we need to choose independently drawn unbiased pattern components $\xi_i^\mu \in \{-1,1\}$ (with equal probabilities) and independently drawn thresholds $\theta_i$ (from some probability distribution $W(\theta)$). Afterwards, by solving equations (9) and (10) in terms of the flow and difussion we can construct the probability distribution for the fluctuations of the dynamical order parameters (8) around mean field trajectories given by Eq. (2). These stochastic fluctuations depend on the actual realization of stored patterns, through the frozen correction term in Eq. (5) and their size is of order $\mathcal{O}\left(N^{-1/2}\right)$.

## 3. Statistical equilibrium and stationary processes

A physical system is said to be in a stationary state when the correlations of the stochastic moments of $q_\mu$, become independent of time translations, that is if

$$\langle q_\mu(t_1+\tau)\,q_\mu(t_2+\tau)\ldots q_\mu(t_n+\tau)\rangle = \langle q_\mu(t_1)\;q_\mu(t_2)\ldots q_\mu(t_n)\rangle$$

for any time $\tau > 0$, and $\mu = 1,\ldots p$ [21]; this condition is satisfied in particular by any system in thermodynamical equilibrium. However, thermodynamical equilibrium requires aditionally the presence of detailed balance, which means that at microscopical level the probability for a 'direct' process is equal to the probability of the 'reverse' process, so there is no net probability current between any two microscopic states of the system. The equilibrium state has strong stability properties: any perturbation acting on it will die out fast; this behaviour is reflected in the second law of thermodynamics, according to which equilibrium is reached in an irreversible fashion and corresponds to a maximum entropy at constant energy, or minimum free energy at constant temperature. However, all this scheme is no longer useful for systems with non-symmetric interactions, as in this case it is not possible to define an energy function and detailed balance no longer holds, so it is not possible to talk about equilibrium. However, it is observed that these kind of systems still evolves from less probable to more probable states, and this evolution leads them towards more privileged parts of the 'space of states'. Therefore, for a system not satisfying detailed balance we can define an "extended entropy" [21] as follows:

$$S(t) = -H(t) \quad \text{with} \quad H(t) \equiv \int d\boldsymbol{q}\,\mathcal{P}_t(\boldsymbol{q}) \ln \mathcal{P}_t(\boldsymbol{q}) \quad (11)$$

(with Boltzmann's constant $k_B \equiv 1$). This extended entropy is an extensive quantity, which, according to the H-theorem, increases monotonically until it reaches its maximum value, provided it exists a stationary state; and this result is valid for any system satisfying, or not, detailed balance. The contribution to the extended entropy, due to finite size fluctuations, at time $t$ is obtained by substituting (8) into the previous expression (11). In this way we find

$$S_q(t) = -\langle \ln \mathcal{P}_t(\mathbf{q}) \rangle_{\mathbf{q};t} = \frac{1}{2} \ln \{\det [\boldsymbol{\Xi}(t)]\} + C,$$

where C is a constant. Therefore, the increase in entropy between two instants $t_1$ and $t_2$ such that $t_1 < t_2$, is given by

$$\Delta S = \ln \left\{ \frac{\det [\boldsymbol{\Xi}(t_2)]}{\det [\boldsymbol{\Xi}(t_1)]} \right\}^{\frac{1}{2}} \geq 0,$$

which implies $\det [\boldsymbol{\Xi}(t_2)] \geq \det [\boldsymbol{\Xi}(t_1)]$. If we combine the Wronski identity $\frac{d}{dt} \ln \{\det [\boldsymbol{B}]\} = Tr \left[\boldsymbol{B}^{-1} \frac{d}{dt} \boldsymbol{B}\right]$ with Eq. (10) describing the time evolution of the second moments, we find

$$\frac{d}{dt} \{\ln [\det \, \boldsymbol{\Xi}(t)]\} = 2\, \mathrm{Tr} \left[\boldsymbol{D}\, \boldsymbol{\Xi}^{-1} - \boldsymbol{L}\right],$$

which implies that in any relaxation process it is satisfied

$$\mathrm{Tr} \left[\boldsymbol{D}\, \boldsymbol{\Xi}^{-1}\right] \geq \mathrm{Tr} [\boldsymbol{L}],$$

where the equal sign applies when the system reaches a stationary state. Therefore, Boltzmann's H-theorem tells us that if a stationary state exists, it satisfies the expression $Tr \left[\boldsymbol{D}\, \boldsymbol{\Xi}^{-1}\right] = Tr [\boldsymbol{L}]$. We can compare this result to the stronger condition $\boldsymbol{D}\, \boldsymbol{\Xi}^{-1} = \boldsymbol{L}$, satisfied by symmetric systems in an equilibrium state (therefore, satisfying detailed balance) [16].

## 4. Non detailed balance, rotating probability current and the existence of stationary states

In this section it is given another way to characterize the difference between stationary and equilibrium states, in the framework of NN. We also find the conditions needed for the existence of a stationary state. The first objective is reached if we analyze the probability current in the space of fluctuation variables by rewriting the FPE as a continuity equation for the probability flux $\boldsymbol{J}_t(\boldsymbol{q})$ as follows:

$$\frac{\partial}{\partial t} P_t(\boldsymbol{q}) + \nabla \cdot \boldsymbol{J}_t(\boldsymbol{q}) = 0,$$

with

$$\boldsymbol{J}_t(\boldsymbol{q}) = \boldsymbol{J}_t^{mec}(\boldsymbol{q}) + \boldsymbol{J}_t^{dis}(\boldsymbol{q}), \quad (12)$$

where we have separated the probability flux in two terms, one related to the flux, and the other to the diffusion; for an Ornstein-Uhlenbeck process, like this one, these are given by

$$\boldsymbol{J}_t^{mec}(\boldsymbol{q}) = -\mathcal{P}_t(\boldsymbol{q}) \langle \boldsymbol{F}[\boldsymbol{q}, t] \rangle,$$

$$\boldsymbol{J}_t^{dis}(\boldsymbol{q}) = \mathcal{P}_t(\boldsymbol{q}) D \Xi^{-1} [\boldsymbol{q}(t) - \langle \boldsymbol{q}(t) \rangle]. \quad (13)$$

Van Kampen [22] has noted that the first term, resulting from the presence of an external field, can be called 'purely mechanical' as it has the characteristics of a Liouville flux, while the second term is 'purely diffusive' as it is related to an irreversible diffusive process.

From expression in Eq. (12), it should be clear that when the divergence of the total probability current $\boldsymbol{J}_t(\boldsymbol{q}) = \boldsymbol{J}_t^{mec}(\boldsymbol{q}) + \boldsymbol{J}_t^{dis}(\boldsymbol{q})$ is zero the system is in a stationary state. If we make $d\langle \boldsymbol{q} \rangle / dt = 0$, $d\,\boldsymbol{\Xi}/dt = 0$ in Eqs. (9-10) for long times, we obtain

$$\langle \boldsymbol{F}[\boldsymbol{q}] \rangle = 0 \implies \langle \boldsymbol{q} \rangle_\infty = -L_\infty^{-1} \boldsymbol{K}_\infty \quad (14)$$

$$L_\infty \Xi_\infty + \Xi_\infty L_\infty^\dagger = 2 D_\infty \quad (15)$$

The first of Eqs. (14) shows that, in general, the mechanical probability current vanishes. On the other hand, Eq. (15) is known as the fluctuation-dissipation theorem (FDT), and it shows that the correlation $\Xi_{\mu\nu}$ may be in a stationary state as a consequence of the compensation between damping, represented by the convection matrix $\boldsymbol{L}$ (6) and diffusion, represented by $\boldsymbol{D}$ (7). In this limit, we also have $\boldsymbol{L}_\infty = \boldsymbol{I} - \beta\, \boldsymbol{D}_\infty \boldsymbol{A}$.

This state is an equilibrium state if, additionally, the probability current is irrotational, that is, if $\partial J_\mu / \partial q_\nu = \partial J_\nu / \partial q_\mu$, for all $\mu$, $\nu$. But this happens only for $A_{\mu\nu} = A_{\nu\mu}$, which implies symmetric interactions, as expected. On the contrary, rotational probability current is present when asymmetric interactions appear and detailed balance cannot exist. Besides, $\frac{\partial}{\partial t} P_\infty(\boldsymbol{q}) = 0$ implies that $\nabla \cdot \boldsymbol{J}_t(\boldsymbol{q}) = 0$, so that, for NN with symmetric $\boldsymbol{A}$, there will be detailed balance in the probability current whenever there is a compensation between the mechanical and the dissipative currents, that is:

$$\frac{\partial}{\partial q_\mu} J_\mu = \frac{\partial}{\partial q_\mu} \left(J_\mu^{mec} + J_\mu^{dis}\right) = 0.$$

In particular, the condition $A_{\mu\nu} = A_{\nu\mu}$ is satisfied by the use of the Hebb rule or a variation of it where each pattern $\{\xi^\mu\}$ is given a different weight in the learning prescription [24]. Therefore, divergence and rotational operating on the probability current, $\boldsymbol{J}_t(\boldsymbol{q})$, give us two conditions to characterize stationary and equilibrium states, respectively, in the framework of NN.

To complete this discussion from the formal point of view, we need to present the algebraic conditions for the existence of stationary states. Obviously, the first condition is that the convection matrix should be nonsingular. But in addition, we need to find the formal solution to the autocorrelation matrix $\boldsymbol{\Xi}$. We start from the FDT: $L_\infty\, \Xi_\infty + \Xi_\infty\, L_\infty^\dagger = 2\, D_\infty$ and by choosing the eigenvectors $\{\mid \lambda_n\rangle\}$ of the matrix $L_\infty^\dagger L_\infty$ as the basis to represent our matrixes. Then we get $L_\infty^\dagger L_\infty = \lambda_n \mid \lambda_n\rangle$, and multiplying (15) by $L_\infty^\dagger$ on the left, and by $L_\infty$ on the right, we obtain

$$\langle\lambda_n|L_\infty^\dagger\, L_\infty\, \Xi_\infty\, L_\infty|\lambda_m\rangle + \langle\lambda_n|L_\infty^\dagger\, \Xi_\infty\, L_\infty^\dagger L_\infty|\lambda_m\rangle = \langle\lambda_n|2\, L_\infty^\dagger\, D_\infty\, L_\infty|\lambda_m\rangle.$$

From Eq. (6) we know that if the interaction matrix satisfies $A_{\mu\nu} = \pm\, A_{\nu\mu}$, then the convection matrix $\boldsymbol{L}$ is also symmetric or antisymmetric, respectively. In this case we can use that $\langle\lambda_n \mid L_\infty^\dagger\, L_\infty = \pm\lambda_n\, \langle\lambda_n \mid$, rearranging and multiplying by $L_\infty^{-1}$ from the right we arrive to the solution

$$\Xi_{n,n'}^\infty = \langle\lambda_n \mid L_\infty^\dagger\, D_\infty\, L_\infty\, G_n\, L_\infty^{-1} \mid \lambda_{n'}\rangle$$

with

$$G_n = 2\sum_{m=1}^{p} \frac{\mid\lambda_m\rangle\,\langle\lambda_n\mid}{\pm\lambda_n + \lambda_m}.$$

Then we find, as expected, that if the interaction matrix $A_{\mu\nu}$ is symmetric, then the solution exists. On the other hand, for an antisymmetric interaction matrix it is a condition for having a stationary state that all eigenvalues $\{\lambda_n\}$ should be different. For other types of interactions it is not possible to know, in general, if the solution to the FDT exists (15), so each particular problem must be analyzed separately, as we will do it in the next section by analyzing a system previously studied [16].

## 5. Study of a system with a fixed point but without detailed balance

In this section we will study the dynamics of a particular NN that, even though due to its structure it is unable to satisfy detailed balance, it has a stable fixed point in the limit: $N \to \infty$. In this neural network, $p$ binary patterns $\{\xi_i^\mu\} = \{+1,-1\}^N$, randomly chosen with equal probability for each of their two values, have been stored according to a $(p \times p)$ asymmetric interaction matrix given by $A_{\mu\nu} = \delta_{\mu\nu} + \varepsilon\, \delta_{\mu 1}\delta_{\nu 2}$, so $\varepsilon$ can be seen as a parameter breaking the symmetry. As we will corroborate later, in the infinite version it only matters the probability distribution $P(\xi_i^\mu)$, while in the finite case it is also relevant the particular pattern realization $\{\xi_i^\mu\}$ (measured by the overlap between pairs of patterns). As we mentioned before, the time evolution of the deterministic part of the dynamical order parameter, $m_\mu^*(t)$, can be obtained by numerical solution of the Liouville equation, Eq. (2).

If $\varepsilon = 0$ the characteristics of the system are as follows: there is complete symmetry between all the Liouville equations for the parameters $m_\mu^*$ (2), the infinite system has $2p$ fixed points related to pure states (only one order parameter is different to zero) given by $m_\mu^* = \pm m_\beta$, with $\mu = 1,\ldots,p$ whose values are temperature dependent and are given by the solution of $m_\beta = \tanh(\beta m_\beta)$. Besides, there are other undesirable stable states related to mixtures of the order parameters (more than one of them are simultaneously different to zero), with smaller basins of attraction [25]. The system evolves towards one or another of the fixed points, depending on the initial conditions.

In the case $\varepsilon \neq 0$ the symmetry between patterns is lost, and the shape of the basins and the placing of the attractors depend on the value of $\varepsilon$. The convection matrix is given by

$$L_{\mu\nu} = \delta_{\mu\nu} - \beta\, \langle\xi_\mu\xi_\nu\left[1 - \tanh^2\beta\,(\boldsymbol{\xi}\cdot\boldsymbol{Am} + \theta)\right]\rangle_{\xi,\theta} + \beta\,\varepsilon\,\delta_{\nu 2}\,\langle\xi_\mu\xi_1\left[1 - \tanh^2\beta\,(\boldsymbol{\xi}\cdot\boldsymbol{Am} + \theta)\right]\rangle_{\xi,\theta}$$

and its determinant is different from zero, so the solution to $\langle\boldsymbol{q}\rangle_\infty = -\boldsymbol{L}_\infty^{-1}\boldsymbol{K}_\infty$ exists. For simplicity, we will analyze the stationary states for $p = 2$; as generalization to other values of $p$ is straightforward and does not add new insight into the physics of the problem. In this case, at $T = 0$ for $0 <\mid\varepsilon\mid< 1$ there exist four attractors, two of them related to pattern $\{\boldsymbol{\xi}^1\}$ and two others having a big overlap with pattern $\{\boldsymbol{\xi}^2\}$ and a small, but still macroscopic overlap with pattern $\{\boldsymbol{\xi}^1\}$; for $\mid\varepsilon\mid> 1$ there exist only the two attractors related to pattern '1'. For higher temperatures an analytical expression has been found for $\varepsilon_c(T)$ [16], the critical value of $\varepsilon$ where stability is lost. The FDT (15) can be summarized into three relevant equations to be written as $S\boldsymbol{\Phi} = \boldsymbol{\Lambda}$, with $\boldsymbol{\Phi} = (\Xi_{11}, \Xi_{12}, \Xi_{22})$ and $\boldsymbol{\Lambda} = (D_{11}, D_{12}, D_{22})$. For the NN under consideration, $S$ is nonsingular whenever $(L_{11})^2 = (L_{22})^2 \neq \sqrt{L_{12}L_{21}}$. Since this condition is met, we can conclude that the stationary state exists. Then the FPE can be written as $\nabla\cdot\boldsymbol{J} = 0$ to establish that compensation between damping and diffusion occurs according to the analysis written before.

It is just a matter of algebra to study the final size effects close to the stationary state $\boldsymbol{m}_\infty^* = m_\beta\,(1,0)$. According to Eq. (9), at the stationary state, $m_\mu^* = m_\beta\,\delta_{\mu 1}$, the average flux $\langle\boldsymbol{F}\rangle_\infty$ is equal to zero and the value of the average fluctuations $\langle\mathbf{q}\rangle_\infty$ depends on the explicit configuration of patterns and it is weighted by the overlap between patterns: $R_{12} = (1/\sqrt{N})\sum_i \xi_i^1\xi_i^2$.

The fixed point $\boldsymbol{m}_\infty^*$ is an stable attractor when the overlap between patterns is non negative ($R_{12} \geq 0$), while for the case $R_{12} < 0$, the state of the system escapes from this basin of attraction and evolves towards $m^* = m_\beta(-1,0)$ [4]; this is an important finite size effect, since escaping is impossible in the infinite case ($R \equiv 0$). The autocorrelation matrix is given by the next expression

$$\Xi_{\mu\nu} = H_m\left(T\right)\left\{\delta_{\mu\nu}\left[1+\delta_{\mu 1}\delta_{\nu 1}\,\frac{1}{2}\beta^2\varepsilon^2 H_m^2\left(T\right)\right] + \frac{1}{2}\,\beta\varepsilon H_m\left(T\right)\left[\delta_{\mu 1}\delta_{\nu 2}+\delta_{\mu 2}\delta_{\nu 1}\right]\right\},$$

in which

$$H_m(T) = \frac{1-(\boldsymbol{m}^*)^2}{1-\beta[1-(\boldsymbol{m}^*)^2]},$$
$$H_m(0)=0, \quad H_m(1)=\infty,$$

so the contribution to the extended entropy, due to finite size fluctuations, is given by

$$S_q\left(\infty\right) = \frac{1}{2}\ln\left\{\det\left[\Xi_\infty\right]\right\} + C = H_m\left(T\right)\left[1+\frac{1}{4}\beta^2\varepsilon^2 H_m^2\left(T\right)\right].$$

This expression allows to see that the entropy associated to the stationary state of the system not satisfying detailed balance, $S_{est}$, can be written as follows

$$S_{est} = S_{eq} + \frac{1}{2}\,\ln\left[1+\frac{1}{4}\beta^2\varepsilon^2 H_m^2\left(T\right)\right],$$

where $S_{eq} = (1/2)\ln\left[H_m\left(t\right)\right]+C$. The rotation of the probability current: $\nabla\times\boldsymbol{J}_\infty\left(\boldsymbol{q}\right) = -\mathcal{P}_\infty\left(\mathbf{q}\right)\beta\varepsilon\left(1-m_\beta^2\right)$, produces this additional entropy.

## 6. Discussion and conclusions

In general terms, in the scientific literature there exists some confusion about the terms ”stationarity” and ”equilibrium” and sometimes they are even treated as synonymous. In this work we discussed the conceptual differences between these two terms from a formal point of view and provided a set of tools that can be used to illustrate this difference in terms of the analysis of finite size effects. For this purpose it was introduced the concept of Extended Entropy *à la* Boltzmann, and it was shown that in any process leading towards a stationary state it is satisfied the relationship Tr $\left[\boldsymbol{D}\,\boldsymbol{\Xi}^{-1}\right] \geq$ Tr $[\boldsymbol{L}]$. This condition is less restrictive than the condition satisfied by a process leading to an equilibrium state, namely $\boldsymbol{D}\,\boldsymbol{\Xi}^{-1} \geq \boldsymbol{L}$, which can only be reached by symmetric systems satisfying detailed balance. We also showed how the lack of detailed balance is related to the rotation of the probability current for the finite size fluctuations $\boldsymbol{q}(t)$. For the case of antisymmetric NN, which are often used to retrieve patterns sequentially [12], we demonstrated that they can reach a stationary state only if its convection matrix is nonsingular and the eigenvalues of $\{\mid\lambda_n\rangle\}$ are all different.

We demonstrated our results by examining the behaviour of a peculiar system whose infinite version has fixed points even though these states do not satisfy detailed balance, and therefore, it cannot reach an equilibrium state. We demonstrated how, if we consider finite size effects, these ”fixed points” transform into the wider category of ”attractors” of the dynamics which keep the system moving around but never reaching them, as the probability distribution of the finite size fluctuations has a rotational different from zero.

Our theory describes finite size effects around the order parameter values which are exact in the thermodynamical limit, and therefore help us to find out how the finite size of the sample modifies the macroscopic behaviour of the system. These fluctuations, originated by the finite size of the system, are specific for each particular system and depend on the specific set of stored patterns.

As we mentioned before, this theory can be applied in principle, to study finite size effects in any kind of finite, separable recurrent NN with sequential dynamics and away from saturation. However, we should keep in mind that this theory is correct only in the low storage regime as in the derivation of the theory it is essential the assumption that the overlap between any pair of stored patterns (chosen at random) goes to zero in the thermodynamical limit. In the case of systems with parallel updating a theory to describe finite size effects has not yet been developed.

## Acknowledgments

This work was partially supported by grant DGAPA IN123098 (UNAM).

1. A. C. C. Coolen and D. Sherrington, in *Mathematical Approaches to Neural Networks*, ed. Taylor JG, (North-Holland, Amsterdam, 1993) 293.
2. B. M. Forrest, *J. Phys.* **A21** (1988) 245.
3. G. A. Kohring, *J. Phys. A: Math. Gen.* **23** (1990) 2237.
4. L. Viana, A. Castellanos and A. A. C. Coolen, *Lect. Notes in Comp. Sci.* **1606** (1999) 392.
5. E. Domany, J. L. van Hemmen, and K. Schulten, (ed), *Models of Neural Networks II*, (Berlin: Springer, 1994).
6. A. Canning and E. Gardner *J. Phys. A: Math. Gen.* **21** (1988) 3275.
7. A. J. Noest *Europhys. Lett.* **6** (1988) 469.
8. E. Domany, W. Kinzel and R. Meir *J. Phys. A: Math. Gen.* **22** (1989) 2081.

9. A. C. C. Coolen and L. Viana *J. Phys. A: Math. Gen*. **26** 1996 2549-66
10. N. S. Skantzos and A. C. C. Coolen, *J. Phys. A: Math. Gen.* **33** (2000) 1841.
11. N. S. Skantzos and A. C. C. Coolen, *J. Phys. A: Math. Gen.* **33** (2000) 5785.
12. H. Nishimori, T. Nakamura and M. Shiino, *Phys. Rev.* **A41** (1990) 3346.
13. P. Peretto, *J. Phys. France* **49** (1988) 711.
14. J. M. G. Amaro, and J. F. Pérez, *Europhys. Lett.* **5** (1988) 277.
15. A. Castellanos, PhD Thesis CICESE, México (1998).
16. A. Castellanos, A.C.C. Coolen, L. Viana *J. Phys.* **A31** (1998) 6615.
17. A. C. C. Coolen and Th. W. Ruijgrok, *Phys. Rev. A.* **38** (1988) 4253.
18. I. J. Matus, and P. Pérez, *Phys. Rev.* **A41** (1990) 7013.
19. F. Zertuche, R. López-Peña and H. Waelbroeck, *J. Phys.* **A27** (1994) 5879.
20. O. Bernier, *Europhys. Lett.* **16** (1991) 531.
21. N. G. Van Kampen, *Can. J. Phys.* **39** (1961) 551; *Adv. Chem. Phys.* **34** (1976) 245.
22. N. G. Van Kampen, ”Stochastic Processes in Physics and Chemistry” (2nd. ed. North-Holland 1992), p 285-286.
23. According to Pawula’s theorem, a consistent probabilistic theory is obtained in only two cases a) if only the first non trivial order in $(1/N)$ is kept, or b) if it is maintained the whole Kramers-Moyal expansion.
24. L. Viana, *J. Phys.* **49** (1988) 167.
25. L. Viana and A. C. C. Coolen, *J. Phys.* **I 3** (1993) 777.