C. R. Acad. Sci. Paris, Ser. I 352 (2014) 841–846

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Dynamical systems/Probability theory

Asymptotic description of stochastic neural networks. I. Existence of a large deviation principle

Description asymptotique de réseaux de neurones stochastiques. I. Existence d'un principe de grandes déviations

Olivier Faugeras, James Maclaurin

Inria Sophia-Antipolis Méditerranée, NeuroMathComp Group, France

A R T I C L E   I N F O

Article history:
Received 28 April 2014
Accepted after revision 26 August 2014
Available online 23 September 2014

Presented by the Editorial Board

A B S T R A C T

We study the asymptotic law of a network of interacting neurons when the number of neurons becomes infinite. The dynamics of the neurons is described by a set of stochastic differential equations in discrete time. The neurons interact through the synaptic weights that are Gaussian correlated random variables. We describe the asymptotic law of the network when the number of neurons goes to infinity. Unlike previous works which made the biologically unrealistic assumption that the weights were i.i.d. random variables, we assume that they are correlated. We introduce the process-level empirical measure of the trajectories of the solutions into the equations of the finite network of neurons and the averaged law (with respect to the synaptic weights) of the trajectories of the solutions into the equations of the network of neurons. The result (Theorem 3.1 below) is that the image law through the empirical measure satisfies a large deviation principle with a good rate function. We provide an analytical expression of this rate function in terms of the spectral representation of certain Gaussian processes.

R É S U M É

Nous considérons un réseau de neurones décrit par un système d'équations différentielles stochastiques en temps discret. Les neurones interagissent au travers de poids synaptiques qui sont des variables aléatoires gaussiennes corrélées. Nous caractérisons la loi asymptotique de ce réseau lorsque le nombre de neurones tend vers l'infini. Tous les travaux précédents faisaient l'hypothèse, irréaliste du point de vue de la biologie, de poids indépendants. Nous introduisons la mesure empirique sur l'espace des trajectoires solutions des équations du réseau de neurones de taille finie et la loi moyennée (par rapport aux poids synaptiques) des trajectoires de ces solutions. Le résultat (théorème 3.1 ci-dessous) est que l'image de cette loi par la mesure empirique satisfait un principe de grandes déviations avec une bonne fonction de taux, dont nous donnons une expression analytique en fonction de la représentation spectrale de certains processus gaussiens.

E-mail addresses: olivier.faugeras@inria.fr (O. Faugeras), james.maclaurin@inria.fr (J. Maclaurin).

http://dx.doi.org/10.1016/j.crma.2014.08.018
1631-073X/

Version française abrégée

Nous considérons le problème de la description de la dynamique asymptotique d'un ensemble de $2n+1$ neurones lorsque ce nombre tend vers l'infini. Ce problème est motivé par un désir de parcimonie dans la description, par celui de rendre compte de l'apparition de phénomènes émergents, ainsi que par celui de comprendre les effets de taille finie. Nous considérons donc un réseau de $2n+1$ neurones interconnectés, dont la dynamique commune (en temps discret) obéit aux équations stochastiques (1). Dans celles-ci apparaissent les poids synaptiques ou coefficients de couplage notés J^n_{ij}, qui sont des variables aléatoires gaussiennes corrélées. Pour répondre à la question posée, nous considérons la loi, notée Q^{V_n}, de la solution à (1) moyennée par rapport aux poids synaptiques ou, plus précisément, l'image Π^n de cette loi par la mesure empirique (3). Nous montrons dans le théorème 3.1 que cette loi satisfait un principe de grandes déviations avec une bonne fonction de taux H, dont nous donnons une expression analytique dans la définition 3.1 et les équations (9) et (12). Ce travail généralise au cas des poids synaptiques corrélés celui d'auteurs comme Sompolinsky [11] et Moynot et Samuelides [8], qui ont considéré le cas de poids synaptiques indépendants. Dans ce cas, plus simple d'un point de vue mathématique, mais beaucoup moins réaliste d'un point de vue biologique, on observe le phénomène de propagation du chaos. Nous montrons dans un second article [5] que la bonne fonction de taux a un minimum unique, que nous caractérisons complètement. La propagation du chaos n'a pas lieu, mais la représentation est parcimonieuse dans un sens défini dans [5].

1. Introduction

1.1. Neural networks

Our goal is to study the asymptotic behaviour and large deviations of a network of interacting neurons when the number of neurons becomes infinite. A more detailed exposition of this work, with proofs, may be found in [4].

Sompolinsky successfully explored this particular topic [11] for fully connected networks of neurons. In his study of the continuous time dynamics of networks of rate neurons, Sompolinsky and his colleagues assumed that the synaptic weights, were i.i.d. random variables with zero mean Gaussian laws. The main result they obtained (using the local chaos hypothesis) under the previous hypotheses is that the averaged law of the neurons dynamics is chaotic in the sense that the averaged law of a finite number of neurons converges to a product measure as the system gets very large.

The next efforts in the direction of understanding the averaged law of neurons are those of Cessac, Moynot, and Samuelides [1,7,8,2,10]. From the technical viewpoint, the study of the collective dynamics is done in discrete time. Moynot and Samuelides obtained a large deviation principle and were able to describe in detail the limit averaged law that had been obtained by Cessac using the local chaos hypothesis and to prove rigorously the propagation of chaos property.

One of the next outstanding challenges is to incorporate in the network model the fact that the synaptic weights are not independent and in effect, according to experimentalists, often highly correlated. Our problem thus resembles that of a random walk in a mixing random environment [12,9].

The problem whose solution we announce in this paper and in [5] is the following. Given a completely connected network of neurons in which the synaptic weights are Gaussian correlated random variables, can we describe the asymptotic law of the network when the number of neurons goes to infinity?

1.2. Mathematical framework

For some positive integer $n \geq 0$, we let $V_n = \{j \in \mathbb{Z} : |j| \leq n\}$, and $|V_n| = 2n+1$. The finite-size neural network below is indexed by points in V_n. We work in discrete time, over times $t \in \{0, 1, \ldots, T\}$, for some positive integer T. The state variable for each neuron is in $\mathbb{R}$, and the path space is $\mathcal{T} = \mathbb{R}^{T+1}$. We equip $\mathcal{T}$ with the Euclidean topology, $\mathcal{T}^{\mathbb{Z}}$ with the cylindrical topology, and denote the Borelian σ-algebra generated by this topology by $\mathcal{B}(\mathcal{T}^{\mathbb{Z}})$.

The equation describing the time variation of the membrane potential U^j of the jth neuron writes:

$$U^j_t = \gamma U^j_{t-1} + \sum_{i \in V_n} J^n_{ji} f\left(U^i_{t-1}\right) + \theta^j + B^j_{t-1}, \qquad U^j_0 = u^j_0, \quad j \in V_n,\ t = 1, \ldots, T, \tag{1}$$

$f : \mathbb{R} \to]0, 1[$ is a monotonically increasing Lipschitz continuous bijection. γ is in $[0, 1)$ and determines the time scale of the intrinsic dynamics of the neurons. The B^j_ts are i.i.d. Gaussian random variables distributed as $\mathcal{N}_1(0, \sigma^2)$.[1] They represent the fluctuations of the neurons' membrane potentials. The θ^js are i.i.d. as $\mathcal{N}_1(\bar{\theta}, \theta^2)$. They are independent of the B^i_ts and represent the current injected in the neurons. The u^j_0s are i.i.d. random variables, each governed by the law μ_I.

The J^n_{ij}s are the synaptic weights. J^n_{ij} represents the strength with which the 'presynaptic' neuron j influences the 'postsynaptic' neuron i. They arise from a stationary Gaussian random field specified by its mean and covariance function

$$\mathbb{E}\left[J^n_{ij}\right] = \frac{\bar{J}}{|V_n|}, \qquad \operatorname{cov}\left(J^n_{ij} J^n_{kl}\right) = \frac{1}{|V_n|} \Lambda\left((k-i) \bmod V_n, (l-j) \bmod V_n\right),$$

[1] We note $\mathcal{N}_p(m, \Sigma)$ the law of the p-dimensional Gaussian variable with mean m and covariance matrix Σ.

Λ is positive definite, let $\tilde{\Lambda}$ be the corresponding (positive) Fourier transform. We make the technical assumption that the summation over both indices of the series $(\Lambda(i,j))_{i,j\in\mathbb{Z}}$ is absolutely convergent to $\Lambda^{\mathrm{sum}} > 0$. We write $\Lambda^{\mathrm{sum}}_{\mathrm{min}} = \inf_{n\geq 0}\sum_{j,k\in V_n}\Lambda(j,k)$ and assume that $\Lambda^{\mathrm{sum}}_{\mathrm{min}} > 0$.

We note J^n the $|V_n|\times|V_n|$ matrix of the synaptic weights, $J^n = (J^n_{ij})_{i,j\in V_n}$.

The process (Y^j) defined by

$$Y^j_t = \gamma Y^j_{t-1} + \bar{\theta} + B^j_{t-1}, \quad j\in V_n,\ t = 1,\cdots,T, \qquad Y^j_0 = u^j_0$$

is stationary and independent. The law of each Y^j is easily found to be given by $P = (\mathcal{N}_T(0_T, \sigma^2 \mathrm{Id}_T)\otimes\mu_I)\circ\Psi$, where $\Psi : \mathcal{T}\to\mathcal{T}$ is the following affine bijection. The joint law of (Y^k) (for $k\in V_n$) is written as $P^{\otimes V_n}$, and the joint law of all (Y^j) is written as $P^{\mathbb{Z}}$. Writing $v = \Psi(u)$, we define:

$$\begin{cases} v_0 = \Psi_0(u) = u_0, \\ v_s = \Psi_s(u) = u_s - \gamma u_{s-1} - \bar{\theta} \quad s = 1,\cdots,T. \end{cases} \tag{2}$$

We extend Ψ to a mapping $\mathcal{T}^{\mathbb{Z}}\to\mathcal{T}^{\mathbb{Z}}$ componentwise. We now introduce some more notation.

For some topological space Ω equipped with its Borelian σ-algebra $\mathcal{B}(\Omega)$, we denote the set of all probability measures by $\mathcal{M}(\Omega)$. We equip $\mathcal{M}(\Omega)$ with the topology of weak convergence. For some $\mu\in\mathcal{M}(\mathcal{T}^{\mathbb{Z}})$ governing a process $(X^j)_{j\in\mathbb{Z}}$, we let $\mu^{V_n}\in\mathcal{M}(\mathcal{T}^{V_n})$ denote the marginal governing $(X^j)_{j\in V_n}$. For some $X\in\mathcal{T}$ and $0\leq a\leq b\leq T$, $X_{a,b}$ denotes the $b-a+1$-dimensional subvector of X. We let $\mu_{a,b}\in\mathcal{M}(\mathcal{T}^{\mathbb{Z}}_{a,b})$ denote the marginal governing $(X^j_{a,b})_{j\in\mathbb{Z}}$. For some $j\in\mathbb{Z}$, let the shift operator $\mathcal{S}^j : \mathcal{T}^{\mathbb{Z}}\to\mathcal{T}^{\mathbb{Z}}$ be $S(\omega)^k = \omega^{j+k}$. We let $\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$ be the set of all stationary probability measures μ on $(\mathcal{T}^{\mathbb{Z}}, \mathcal{B}(\mathcal{T}^{\mathbb{Z}}))$ such that for all $j\in\mathbb{Z}$, $\mu\circ(\mathcal{S}^j)^{-1} = \mu$.

We next introduce the following definitions.

Definition 1.1. For each measure $\mu\in\mathcal{M}(\mathcal{T}^{V_n})$ or $\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$ we define $\underline{\mu}$ to be $\mu\circ\Psi^{-1}$.

Definition 1.2. Let $\mathcal{E}_2$ be the subset of $\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$ defined by $\mathcal{E}_2 = \{\mu\in\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}}) \mid \mathbb{E}^{\underline{\mu}_{1,T}}[\|v^0\|^2] < \infty\}$. Let $\mathbb{M}^+_T$ be the set of all $T\times T$ symmetric matrices with nonnegative eigenvalues. For $\mu\in\mathcal{E}_2$, we have by Herglotz' theorem that there exists an $\mathbb{M}^+_T$-valued spectral measure $\tilde{v}^{\mu}$ on $([-\pi,\pi], \mathcal{B}([-\pi,\pi]))$ such that for all $j,k\in\mathbb{Z}$,[2]

$$E^{\underline{\mu}_{1,T}}\left[v^{j\dagger}v^k\right] = \frac{1}{2\pi}\int\limits_{[-\pi,\pi]}\exp\bigl(\mathrm{i}\theta(k-j)\bigr)\,\mathrm{d}\tilde{v}^{\mu}(\theta).$$

Let $p_n : \mathcal{T}^{V_n}\to\mathcal{T}^{\mathbb{Z}}$ be such that $p_n(\omega)^k = \omega^{k \bmod V_n}$. Here, and throughout the paper, we take $k \bmod V_n$ to be the element $l\in V_n$ such that $l = k \bmod |V_n|$. Define the process-level empirical measure $\hat{\mu}_n : \mathcal{T}^{V_n}\to\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$ as

$$\hat{\mu}_n(\omega) = \frac{1}{|V_n|}\sum_{k\in V_n}\delta_{S^k p_n(\omega)}. \tag{3}$$

We define the process-level entropy to be, for $\mu\in\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$, $I^{(3)}(\mu, P^{\mathbb{Z}}) = \lim_{n\to\infty}\frac{1}{|V_n|}I^{(2)}(\mu^{V_n}, P^{\otimes V_n})$. If $\mu\notin\mathcal{E}_2$, then $I^{(3)}(\mu, P^{\mathbb{Z}}) = \infty$. Here $I^{(2)}$ is the *relative entropy*. For further discussion, a definition of $I^{(2)}$ and a proof that $I^{(3)}$ is well-defined, see [3].

We note $Q^{V_n}(J^n)$ the element of $\mathcal{M}(\mathcal{T}^{V_n})$, which is the law of the solution to (1) conditioned on J^n. We let $Q^{V_n} = \mathbb{E}^J[Q^{V_n}(J^n)]$ be the law averaged with respect to the weights. The reason for this is that we want to study the empirical measure $\hat{\mu}_n$ on the path space $\mathcal{T}^{\mathbb{Z}}$. There is no reason for this to be a simple problem, since for a fixed interaction J^n, the variables $(U^j)_{j\in V_n}$ are not exchangeable. So we first study the law of $\hat{\mu}_n$ averaged over the interactions.

Finally, we introduce the image laws in terms of which the principal results of this paper are formulated.

Definition 1.3. Let Π^n and R^n in $\mathcal{M}(\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}}))$ be the image laws of Q^{V_n} and $P^{\otimes V_n}$ through the function $\hat{\mu}_n : \mathcal{T}^{V_n}\to\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$ defined by (3):

$$\Pi^n = Q^{V_n}\circ\hat{\mu}_n^{-1} \qquad R^n = P^{\otimes V_n}\circ\hat{\mu}_n^{-1}.$$

[2] We note † the transpose of a vector or matrix.

2. The good rate function

We obtain an LDP for the process with correlations (Π^n) via the (simpler) process without correlations (R^n). To do this we obtain an expression for the Radon–Nikodym derivative of Π^n with respect to R^n. This is done in Propositions 2.4 and 2.5. In Eq. (13) there appear certain Gaussian random variables defined from the right-hand side of the equations of the neuronal dynamics (1). Applying the Gaussian calculus to this expression, we obtain Eq. (14), which expresses the Radon–Nikodym derivative as a function (depending on n) of the empirical measure (3). Using the fact that this function is measurable, we obtain Eq. (15). This equation is essential in a) finding the expression for the function Γ that appears in the rate function H of Definition 3.1, b) proving the lower-bound for Π^n on the open sets, c) proving that the sequence (Π^n) is exponentially tight, and d) proving the upper bound on the compact sets.

The key idea is to associate with every stationary measure μ a certain stationary Gaussian process G^μ, or equivalently a certain Gaussian measure defined by its mean c^μ and its covariance operator K^μ. This allows us to write the Radon–Nikodym derivative as a function of the empirical measure, through writing it as a function of $G^{\hat{\mu}_n}$.

Given μ in $\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$, we define a stationary Gaussian process G^μ, governed by a measure $\mathcal{Q}^\mu \in \mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}}_{1,T})$. For all i, the mean of $G^{\mu,i}_t$ is given by c^μ_t, where

$$c^\mu_t = \bar{J} \int_{\mathcal{T}^{\mathbb{Z}}} f\left(u^i_{t-1}\right) \mathrm{d}\mu(u), \quad t = 1, \cdots, T, i \in \mathbb{Z}. \tag{4}$$

The covariance between the Gaussian vectors $G^{\mu,i}$ and $G^{\mu,i+k}$ is defined to be

$$K^{\mu,k} = \theta^2 \delta_k 1_T{}^\dagger 1_T + \sum_{l=-\infty}^{\infty} \Lambda(k,l) M^{\mu,l}, \tag{5}$$

where 1_T is the T-dimensional vector whose coordinates are all equal to 1 and

$$M^{\mu,k}_{st} = \int_{\mathcal{T}^{\mathbb{Z}}} f\left(u^0_{s-1}\right) f\left(u^k_{t-1}\right) \mathrm{d}\mu(u), \tag{6}$$

The above integrals are well defined because of the definition of f and the fact that the series in (5) is convergent (since the series $(\Lambda(k,l))_{k,\,l\in\mathbb{Z}}$ is absolutely convergent and the elements of $M^{\mu,l}$ are bounded by 1 for all $l \in \mathbb{Z}$). We note $\mathcal{Q}^\mu_{[n]}$ the law of the $|V_n|$-dimensional Gaussian defined by restricting the sum in (5) to $l \in V_n$.

These definitions imply the existence of a Hermitian-valued spectral representation for the sequence $M^{\mu,k}$ (resp. $K^{\mu,k}$) noted $\tilde{M}^\mu$ (resp. $\tilde{K}^\mu$), which satisfies:

$$\tilde{K}^\mu(\theta) = \theta^2 1_T{}^\dagger 1_T + \frac{1}{2\pi} \int_{-\pi}^{\pi} \tilde{\Lambda}(\theta, -\varphi) \tilde{M}^\mu(\mathrm{d}\varphi).$$

Here $\tilde{K}^\mu(\theta) = \sum_{j\in\mathbb{Z}} \exp(-\mathrm{i} j\theta) K^{\mu,j}$, and similarly for $\tilde{M}^\mu$. This allows us to define the spectral representation

$$\tilde{A}^\mu(\theta) = \tilde{K}^\mu(\theta)\left(\sigma^2 \operatorname{Id}_T + \tilde{K}^\mu(\theta)\right)^{-1}. \tag{7}$$

Let $K^{\mu,k}_{[n]}$, $k \in V_n$ be the partial sums of (5) for $l \in V_n$, and $\tilde{K}^\mu_{[n]}(\theta) = \sum_{j\in V_n} \exp(-ij\theta) K^{\mu,j}$. We may thus define $\tilde{A}^\mu_{[n]}(\theta) = \tilde{K}^\mu_{[n]}(\theta)(\sigma^2 \operatorname{Id}_T + \tilde{K}^\mu_{[n]}(\theta))^{-1}$, which in the limit $n \to \infty$ converges to $\tilde{A}^\mu(\theta)$. We write, for $k \in V_n$, $A^{\mu,k}_{[n]} = (|V_n|)^{-1} \sum_{j\in V_n} \exp(2\pi \mathrm{i} jk/|V_n|) \tilde{A}^\mu(2\pi j/|V_n|)$. We next define a functional $\Gamma_{[n]} = \Gamma_{[n],1} + \Gamma_{[n],2}$, which we use to characterise the Radon–Nikodym derivative of Π^n with respect to R^n. Let $\mu \in \mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$ and

$$\Gamma_{[n],1}(\mu) = -\sum_{j\in V_n} \frac{1}{2|V_n|} \log\left(\det\left(\operatorname{Id}_T + \frac{1}{\sigma^2} \tilde{K}^\mu_{[n]}\left(2\pi j/|V_n|\right)\right)\right). \tag{8}$$

Because of the previous remarks, the expression above has a sense. Taking the limit when $n \to \infty$ does not pose any problem, we can define $\Gamma_1(\mu) = \lim_{n\to\infty} \Gamma_{[n],1}(\mu)$. The following lemma, whose proof is straightforward, indicates that this is well defined.

Lemma 2.1. *When n goes to infinity, the limit of* (8) *is given by*

$$\Gamma_1(\mu) = -\frac{1}{4\pi} \int_{-\pi}^{\pi} \log\left(\det\left(\operatorname{Id}_T + \frac{1}{\sigma^2} \tilde{K}^\mu(\theta)\right)\right) \mathrm{d}\theta \quad \textit{for all } \mu \in \mathcal{M}^+_{1,S}(\mathcal{T}^{\mathbb{Z}}). \tag{9}$$

It also follows easily from previous remarks that

Proposition 2.1. *$\Gamma_{[n],1}$ and Γ_1 are bounded below and continuous on $\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$.*

The definition of $\Gamma_{[n],2}(\mu)$ is slightly more technical, but follows naturally from Propositions 2.4 and 2.5. For $\mu \in \mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$ let

$$\Gamma_{[n],2}(\mu) = \int_{\mathcal{T}_{1,T}^{V_n}} \phi^n(\mu, v)\underline{\mu}_{1,T}^{V_n}(\mathrm{d}v), \tag{10}$$

where $\phi^n : \mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}}) \times \mathcal{T}_{1,T}^{V_n} \to \mathbb{R}$ is defined by

$$\phi^n(\mu, v) = \frac{1}{2\sigma^2}\left(\frac{1}{|V_n|}\sum_{j,k\in V_n} {}^{\dagger}\left(v^j - c^\mu\right)A_{[n]}^{\mu,k}\left(v^{k+j} - c^\mu\right) + \frac{2}{|V_n|}\sum_{j\in V_n}\left\langle c^\mu, v^j\right\rangle - \left\|c^\mu\right\|^2\right), \tag{11}$$

where $\langle\cdot,\cdot\rangle$ indicates the usual inner product in $\mathbb{R}^T$. $\Gamma_{[n],2}(\mu)$ is finite in the subset $\mathcal{E}_2$ of $\mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$ defined in Definition 1.2. If $\mu \notin \mathcal{E}_2$, then we set $\Gamma_{[n],2}(\mu) = \infty$.

We define $\Gamma_2(\mu) = \lim_{n\to\infty} \Gamma_{[n],2}(\mu)$. The following proposition indicates that $\Gamma_2(\mu)$ is well defined.

Proposition 2.2. *If the measure μ is in $\mathcal{E}_2$, i.e. if $\mathbb{E}^{\underline{\mu}_{1,T}}[\|v^0\|^2] < \infty$, then $\Gamma_2(\mu)$ is finite and writes:*

$$\Gamma_2(\mu) = \frac{1}{2\sigma^2}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \tilde{A}^\mu(-\theta) : \tilde{v}^\mu(\mathrm{d}\theta) + {}^{\dagger}c^\mu\left(\tilde{A}^\mu(0) - \mathrm{Id}_T\right)c^\mu + 2\mathbb{E}^{\underline{\mu}_{1,T}}\left[{}^t v^0\left(\mathrm{Id}_T - \tilde{A}^\mu(0)\right)c^\mu\right]\right). \tag{12}$$

The ":" symbol indicates the double contraction on the indexes.

It is shown in [4] that $\phi^n(\mu, v)$ defined by (11) is a continuous function of μ that satisfies

$$\phi^n(\mu, v) \geq -\beta_2, \quad \beta_2 = \frac{T\bar{J}^2}{2\sigma^2\Lambda_{\min}^{\mathrm{sum}}}\left(\sigma^2 + \theta^2 + \Lambda^{\mathrm{sum}}\right).$$

By a standard argument, we obtain the following proposition.

Proposition 2.3. *$\Gamma_{[n],2}(\mu)$ is lower-semicontinuous.*

We define $\Gamma_{[n]}(\mu) = \Gamma_{[n],1}(\mu) + \Gamma_{[n],2}(\mu)$. We may conclude from Propositions 2.1 and 2.3 that $\Gamma_{[n]}$ is lower-semicontinuous, hence measurable.

From these definitions, it is relatively easy, and proved in [4], to show that the measure Q^{V_n} is absolutely continuous with respect to $P^{\otimes V_n}$ with a Radon–Nikodym derivative, which can be expressed as a function of the functional $\Gamma_{[n]}$.

Proposition 2.4. *The Radon–Nikodym derivative of Q^{V_n} with respect to $P^{\otimes V_n}$ is given by the following expression:*

$$\frac{\mathrm{d}Q^{V_n}}{\mathrm{d}P^{\otimes V_n}}(u) = \mathbb{E}\left[\exp\left(\frac{1}{\sigma^2}\left(\sum_{j\in V_n}\left\langle \Psi_{1,T}(u^j), G^j\right\rangle - \frac{1}{2}\left\|G^j\right\|^2\right)\right)\right], \tag{13}$$

for all $u \in V_n$, and the expectation being taken against the $2n+1$ T-dimensional Gaussian processes (G^i), $i \in V_n$, given by

$$G_t^i = \sum_{j\in V_n} J_{ij}^n f\left(u_{t-1}^j\right), \quad t = 1, \cdots, T, \quad \text{and the function } \Psi \text{ being defined by (2).}$$

We note that the mean of G_t^i is $c^{\hat{\mu}_n(u)}$ and the covariance $E^{J^n}[G_s^i G_t^j] = K_{[n],st}^{\hat{\mu}_n(u),j-i}$. In fact, the identities above mean that $(G^i)_{i\in V_n}$ converges to $(G^{\mu_e,i})_{i\in\mathbb{Z}}$, where μ_e is defined in [5] and is such that $\hat{\mu}_n \to \mu_e$ almost surely. Using a standard Gaussian calculus, we obtain the following proposition.

Proposition 2.5. *The Radon–Nikodym derivatives write as:*

$$\frac{\mathrm{d}Q^{V_n}}{\mathrm{d}P^{\otimes V_n}}(u) = \exp\left(|V_n|\Gamma_{[n]}\left(\hat{\mu}_n(u)\right)\right), \tag{14}$$

$$\frac{\mathrm{d}\Pi^n}{\mathrm{d}R^n}(\mu) = \exp\left(|V_n|\Gamma_{[n]}(\mu)\right). \tag{15}$$

Here $\mu \in \mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}})$*,* $\Gamma_{[n]}(\mu) = \Gamma_{[n],1}(\mu) + \Gamma_{[n],2}(\mu)$ *and the expressions for* $\Gamma_{[n],1}$ *and* $\Gamma_{[n],2}$ *have been defined in Eqs.* (8) *and* (10)*.*

3. The large deviation principle

We define the function $H : \mathcal{M}_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}}) \to [0, +\infty]$ as follows.

Definition 3.1. Let H be the function $\mathcal{M}^+_{\mathcal{S}}(\mathcal{T}^{\mathbb{Z}}) \to \mathbb{R} \cup \{+\infty\}$ defined by

$$H(\mu) = \begin{cases} +\infty & \text{if } I^{(3)}(\mu, P^{\mathbb{Z}}) = \infty, \\ I^{(3)}(\mu, P^{\mathbb{Z}}) - \Gamma(\mu) & \text{otherwise}, \end{cases}$$

where $\Gamma = \Gamma_1 + \Gamma_2$. We finally state the following theorem.

Theorem 3.1. Π^n *is governed by a large deviation principle with a good rate function* H*.*

The proof is too long to be reproduced here, see [4]. We only give the general strategy. First we prove the lower bound on the open sets. For the upper bound on the closed sets, we simply avoid it by a) proving that (Π^n) is exponentially tight, which allows us to b) restrict the proof of the upper bound to compact sets. The proof of b) is long and technical. It is partially built upon ideas found in [6].

Note that we have found an analytical form for H through Eqs. (9) and (12).

4. Discussion

We have considered the problem of describing the large-size asymptotic dynamics of a neuronal ensemble. We are motivated by the desire for a parsimonious description of such large ensembles, so that we may better recognize emergent phenomena, and also understand finite-size effects. The synaptic weights are correlated: a generalization of the work of Sompolinsky [11] and Moynot and Samuelides [8] where the synaptic weights are independent. In the sequel [5], we characterize the limit μ_e towards which $\hat{\mu}^n$ converges. The correlations in the synaptic weights mean that the limit μ_e does not exhibit the propagation of chaos property, unlike in [11,8], but nevertheless the limit μ_e is a useful sparse representation of the system. The rate function H is another powerful tool which we hope to use in the future to study finite-size effects.

Acknowledgements

This work was supported by INRIA FRM, ERC-NERVI number 227747, European Union Project # FP7-269921 (BrainScales), and Mathemacs # FP7-ICT-2011.9.7

References

[1] B. Cessac, Increase in complexity in random neural networks, J. Phys. I, Fr. 5 (1995) 409–432.
[2] B. Cessac, M. Samuelides, From neuron to neural networks dynamics, Eur. Phys. J. Spec. Top. 142 (2007) 7–88, topics in dynamical neural networks.
[3] J. Deuschel, D. Stroock, H. Zessin, Microcanonical distributions for lattice gases, Commun. Math. Phys. 139 (1991).
[4] O. Faugeras, J. Maclaurin, Asymptotic description of neural networks with correlated synaptic weights, Rapport de recherche RR-8495, INRIA, March 2014.
[5] O. Faugeras, J. Maclaurin, Asymptotic description of stochastic neural networks. II. Characterization of the limit law, C. R. Acad. Sci. Paris, Ser. I 352 (10) (2014) 847–852.
[6] A. Guionnet, Dynamique de Langevin d'un verre de spins, PhD thesis, Université de Paris-Sud, 1995.
[7] O. Moynot, Étude mathématique de la dynamique des réseaux neuronaux aléatoires récurrents, PhD thesis, Université Paul-Sabatier, Toulouse, France, 1999.
[8] O. Moynot, M. Samuelides, Large deviations and mean-field theory for asymmetric random recurrent neural networks, Probab. Theory Relat. Fields 123 (2002) 41–75.
[9] F. Rassoul-Agha, The point of view of the particle on the law of large numbers for random walks in a mixing environment, Ann. Appl. Probab. 31 (3) (2003) 1441–1463.
[10] M. Samuelides, B. Cessac, Random recurrent neural networks, Eur. Phys. J. Spec. Top. 142 (2007) 7–88.
[11] H. Sompolinsky, A. Crisanti, H. Sommers, Chaos in random neural networks, Phys. Rev. Lett. 61 (1988) 259–262.
[12] A. Sznitman, M. Zerner, A law of large numbers for random walks in random environment, Ann. Probab. 27 (1999) 1851–1869.