

A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs

Olivier Faugeras, Jonathan Touboul, Bruno Cessac

▶ **To cite this version:**

Olivier Faugeras, Jonathan Touboul, Bruno Cessac. A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs. [Research Report] RR-6454, INRIA. 2008, pp.58. inria-00258345v3

HAL Id: inria-00258345
https://inria.hal.science/inria-00258345v3

Submitted on 4 Aug 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs

Olivier Faugeras — Jonathan Touboul — Bruno Cessac

N° 6454

August 4, 2008

Thème BIO



ISSN 0249-6399 ISRN INRIA/RR--6454--FR+ENG



A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs

Olivier Faugeras , Jonathan Touboul , Bruno Cessac

Thème BIO — Systèmes biologiques
Projet Odyssée *

Rapport de recherche n° 6454 — August 4, 2008 — 55 pages

Abstract: We deal with the problem of bridging the gap between two scales in neuronal modeling. At the first (microscopic) scale, neurons are considered individually and their behavior described by stochastic differential equations that govern the time variations of their membrane potentials. They are coupled by synaptic connections acting on their resulting activity, a nonlinear function of their membrane potential. At the second (mesoscopic) scale, interacting populations of neurons are described individually by similar equations. The equations describing the dynamical and the stationary mean field behaviors are considered as functional equations on a set of stochastic processes. Using this new point of view allows us to prove that these equations are well-posed on any finite time interval and to provide, by a fixed point method, a constructive method for effectively computing their unique solution. This method is proved to converge to the unique solution and we characterize its complexity and convergence rate. We also provide partial results for the stationary problem on infinite time intervals. These results shed some new light on such neural mass models as the one of Jansen and Rit (Jansen and Rit, 1995): their dynamics appears as a coarse approximation of the much richer dynamics that emerges from our analysis. Our numerical experiments confirm that the framework we propose and the numerical methods we derive from it provide a new and powerful tool for the exploration of neural behaviors at different scales.

Key-words: Mean field analysis, stochastic processes, stochastic differential equations, stochastic networks, stochastic functional equations, random connectivities, multi populations networks, neural mass models

* Odyssée is a joint project between ENS Ulm - INRIA

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

A constructive mean field analysis of multi population neural networks with random synaptic weights and stochastic inputs

Résumé : Nous traitons du problème de combler le fossée entre duex niveaux de modélisation neuronale. A la première échelle (microscopique), les neurones sont considérés individuellement et leur comportement est décrit par des équations différentielles stochastiques qui gouvernent les variations temporelles de leur potentiel de membrane. Ils sont couplés par des connections synaptiques qui agissent sur leur activité, qui est une fonction nonlinéaire de leur potentiel de membrane. A la seconde échelle (mésoscopique) les populations de neurones sont décrites individuellement par le même type d'équation. Les équations qui décrivent les comportements de champ moyen dynamique et stationnaire sont considérées comme des équations fonctionnelles dans un espace de processus stochastiques. Ce nouveau point de vue nous permet de démontrer que ces équations ssont bien posées sur des intervalles de temps finis et de proposer, par une méthode de point fixe, une méthode constructive permettant de calculer efficacement leur unique solution. Nous démontrons que cette méthode converge et caracterisons sa complexité et son taux de convergence. Nous donnons aussi des résultats partiels pour le problème stationnaire sur des intervalles de temps infinis. Ces résultats apportent un nouvel éclairage sur les modéles de masses neuronales tels que celui de Jansen et Rit (Jansen and Rit, 1995): leur dynamique apparaît comme une approximation grossière de la dynamique bien plus riche qui émerge de notre analyse. Nos simulations numériques confirme que le cadre mathématiques que nous proposons et les méthodes numeriques qui en découlent fournissent un outil nouveau et puissant pour l'exploration de comportements neuronaux à différentes échelles.

Mots-clés : Analyse champs moyen, processus stochastiques, équations différentielles stochastiques, réseaux stochastiques, équations stochastiques fonctionnelles, connectivités aléatoires, réseaux multi populations, modèles de masses neuronales

1 Introduction

Modeling neural activity at scales integrating the effect of thousands of neurons is of central importance for several reasons. First, most imaging techniques are not able to measure individual neuron activity ("microscopic" scale), but are instead measuring mesoscopic effects resulting from the activity of several hundreds to several hundreds of thousands of neurons. Second, anatomical data recorded in the cortex reveal the existence of structures, such as the cortical columns, with a diameter of about $50\mu m$ to $1mm$, containing of the order of one hundred to one hundred thousand neurons belonging to a few different species. These columns have specific functions. For example, in the visual cortex V1, they respond to preferential orientations of bar-shaped visual stimuli. In this case, information processing does not occur at the scale of individual neurons but rather corresponds to an activity integrating the collective dynamics of many interacting neurons and resulting in a mesoscopic signal. The description of this collective dynamics requires models which are different from individual neurons models. In particular, if the accurate description of one neuron requires "m" parameters (such as sodium, potassium, calcium conductances, membrane capacitance, etc...), it is not necessarily true that an accurate mesoscopic description of an assembly of N neurons requires Nm parameters. Indeed, when N is large enough averaging effects appear, and the collective dynamics is well described by an effective mean field, summarizing the effect of the interactions of a neuron with the other neurons, and depending on a few effective control parameters. This vision, inherited from statistical physics requires that the space scale be large enough to include a large number of microscopic components (here neurons) and small enough so that the region considered is homogeneous. This is in effect the case of cortical columns.

However, obtaining the equations of evolution of the effective mean field from microscopic dynamics is far from being evident. In simple physical models this can be achieved via the law of large numbers and the central limit theorem, provided that time correlations decrease sufficiently fast. This type of approach has been generalized to such fields as quantum field theory or non equilibrium statistical mechanics. To the best of our knowledge, the idea of applying mean field methods to neural networks dates back to Amari (Amari, 1972; Amari et al, 1977). In his approach, the author uses an assumption that he called the "local chaos hypothesis", reminiscent of Boltzmann's "molecular chaos hypothesis", that postulates the vanishing of individual correlations between neurons, when the number N of neurons tends to infinity. Later on, Crisanti, Sompolinsky and coworkers (Sompolinsky et al, 1988) used a dynamic mean field approach to conjecture the existence of chaos in an homogeneous neural network with random independent synaptic weights. This approach was formerly developed by Sompolinsky and coworkers for spin-glasses (Crisanti and Sompolinsky, 1987a,b; Sompolinsky and Zippelius, 1982), where complex effects such as aging or coexistence of a diverging number of metastable states, renders the mean field analysis delicate in the long time limit (Houghton et al, 1983).

On the opposite, these effects do not appear in the neural network considered in (Sompolinsky et al, 1988) because the synaptic weights are independent (Cessac, 1995) (and especially non symmetric, in opposition to spin glasses). In this case, the Amari approach and the dynamic mean field approach lead to the same mean field equations. Later on, the mean field equations derived by Sompolinsky and Zippelius (Sompolinsky and Zippelius, 1982) for spin-glasses were rigorously obtained by

Ben Arous and Guionnet (Ben-Arous and Guionnet, 1995, 1997; Guionnet, 1997). The application of their method to a discrete time version of the neural network considered in (Sompolinsky et al, 1988) and in (Molgedey et al, 1992) was done by Moynot and Samuelides (Moynot and Samuelides, 2002).

Mean field methods are often used in the neural network community but there are only a few rigorous results using the dynamic mean field method. The main advantage of dynamic mean field techniques is that they allow one to consider neural networks where synaptic weights are random (and independent). The mean field approach allows one to state general and generic results about the dynamics as a function of the statistical parameters controlling the probability distribution of the synaptic weights (Samuelides and Cessac, 2007). It does not only provide the evolution of the mean activity of the network but, because it is an equation on the law of the mean field, it also provides informations on the fluctuations around the mean and their correlations. These correlations are of crucial importance as revealed in the paper by Sompolinsky and coworkers (Sompolinsky et al, 1988). Indeed, in their work, the analysis of correlations allows them to discriminate between two distinct regimes: a dynamics with a stable fixed point and a chaotic dynamics, while the mean is identically zero in the two regimes.

However, this approach has also several drawbacks explaining why it is so seldom used. First, this method uses a generating function approach that requires heavy computations and some "art" for obtaining the mean field equations. Second, it is hard to generalize to models including several populations. Their approach consists in considering that dynamic mean field equations characterize *in fine* a stationary process. It is then natural to search for stationary solutions. This considerably simplifies the dynamic mean field equations by reducing them to a set of differential equations (see section 5) but the price to pay is the unavoidable occurrence in the equations of a non free parameter, the initial condition, that can only be characterized through the investigation of the non stationary case. Hence it is not clear whether such a stationary solution exists, and, if it is the case, how to characterize it. To the best of our knowledge, this difficult question has only been investigated for neural networks in one paper by Crisanti and coworkers (Crisanti et al, 1990).

Different alternative approaches have been used to get a mean field description of a given neural network and to find its solutions. In the neuroscience community, a static mean field study of multi population network activity was developed by Treves in (Treves, 1993). This author did not consider external inputs but incorporated dynamical synaptic currents and adaptation effects. His analysis was completed in (Abbott and Van Vreeswijk, 1993), where the authors considered a unique population of nonlinear oscillators subject to a noisy input current. They proved, using a stationary Fokker-Planck formalism, the stability of an asynchronous state in the network. Later on, Gerstner in (Gerstner, 1995) built a new approach to characterize the mean field dynamics for the Spike Response Model, via the introduction of suitable kernels propagating the collective activity of a neural population in time.

Brunel and Hakim considered a network composed of integrate-and-fire neurons connected with constant synaptic weights (Brunel and Hakim, 1999). In the case of sparse connectivity, stationarity, and considering a regime where individual neurons emit spikes at low rate, they were able to study analytically the dynamics of the network and to show that the network exhibited a sharp transition between a stationary regime and a regime of fast collective oscillations weakly synchronized. Their

approach was based on a perturbative analysis of the Fokker-Planck equation. A similar formalism was used in (Mattia and Del Giudice, 2002) which, when complemented with self-consistency equations, resulted in the dynamical description of the mean field equations of the network, and was extended to a multi population network.

In the present paper, we investigate this question using a new and rigorous approach based on stochastic analysis.

The article is organized as follows. In section 2 we derive from first principles the equations relating the membrane potential of each of a set of neurons as function of the external injected current and noise and of the shapes and intensities of the postsynaptic potentials in the case where these shapes depend only on the post-synaptic neuron (the so-called voltage-based model) and in the case where they depend only on the nature of the presynaptic neurons (the so-called activity-based model). Assuming that the shapes of the postsynaptic potentials can be described by linear (possibly time-dependent) differential equations we express the dynamics of the neurons as a set of stochastic differential equations and give sufficient conditions for the equivalence of the voltage- and activity based descriptions. This allows us to obtain the mean field equations when the neurons belong to P populations whose sizes grow to infinity and the intensities of the postsynaptic potentials are independent Gaussian random variables whose law depend on the populations of the pre- and post-synaptic neurons and not on the individual neurons themselves. These equations can be derived in several ways, either heuristically as in the work of Amari (Amari, 1972; Amari et al, 1977), Sompolinsky (Crisanti et al, 1990; Sompolinsky et al, 1988), and Cessac (Cessac, 1995; Samuelides and Cessac, 2007), or rigorously as in the work of Benarous and Guionnet (Ben-Arous and Guionnet, 1995, 1997; Guionnet, 1997). Our purpose in this article is not their derivation but to prove that they are well-posed and to provide an algorithm for computing their solution. Before we do this we provide the reader with two important examples of such mean field equations. The first example is what we call the simple model, a straightforward generalization of the case studied by Amari and Sompolinsky. The second example is a neuronal assembly model, or neural mass model, as introduced by Freeman (Freeman, 1975) and examplified in Jansen and Rit's cortical column model (Jansen and Rit, 1995).

In section 3 we consider the problem of solutions over a finite time interval $[t_0, T]$. We prove, under some mild assumptions, the existence and uniqueness of a solution of the dynamic mean field equations given an initial condition at time t_0. The proof consists in showing that a nonlinear equation defined on the set of multidimensional Gaussian random processes defined on $[t_0, T]$ has a fixed point. We extend this proof in section 4 to the case of stationary solutions over the time interval $[-\infty, T]$ for the simple model. Both proofs are constructive and provide an algorithm for computing numerically the solutions of the mean field equations.

We then study in section 5 the complexity and the convergence rate of this algorithm and put it to good use: We first compare our numerical results to the theoretical results of Sompolinsky and coworkers (Crisanti et al, 1990; Sompolinsky et al, 1988). We then provide an example of numerical experiments in the case of two populations of neurons where the role of the mean field fluctuations is emphasized.

Along the paper we introduce several constants. To help the reader we have collected in table 1 the most important ones and the place where they are defined in the text.

2 Mean field equations for multi-populations neural network models

In this section we introduce the classical neural mass models and compute the related mean field equations they satisfy in the limit of an infinite number of neurons

2.1 The general model

2.1.1 General framework

We consider a network composed of N neurons indexed by $i \in \{1, \ldots, N\}$ belonging to P populations indexed by $\alpha \in \{1, \ldots, P\}$ such as those shown in figure 1. Let N_α be the number of neurons in population α. We have $N = \sum_{\alpha=1}^{P} N_\alpha$. In the following we are interested in the limit $N \to \infty$. We assume that the proportions of neurons in each population are non-trivial, i.e. :

$$\lim_{N\to\infty} \frac{N_\alpha}{N} = n_\alpha \ \in \ (0,1) \ \forall \alpha \in \{1, \ldots, P\}.$$

If it were not the case the corresponding population would not affect the global behavior of the system, would not contribute to the mean field equation, and could be neglected.

We introduce the function $p : \{1, \ldots, N\} \to \{1, \ldots, P\}$ such that $p(i)$ is the index of the population which the neuron i belongs to.

The following derivation is built after Ermentrout's review (Ermentrout, 1998). We consider that each neuron i is described by its membrane potential $V_i(t)$ or by its instantaneous firing rate $\nu_i(t)$, the relation between the two quantities being of the form $\nu_i(t) = S_i(V_i(t))$ (Dayan and Abbott, 2001; Gerstner and Kistler, 2002), where S_i is sigmoidal.
A single action potential from neuron j is seen as a post-synaptic potential $PSP_{ij}(t-s)$ by neuron i, where s is the time of the spike hitting the synapse and t the time after the spike. We neglect the delays due to the distance travelled down the axon by the spikes.

Assuming that the post-synaptic potentials sum linearly, the average membrane potential of neuron i is

$$V_i(t) = \sum_{j,k} PSP_{ij}(t - t_k),$$

where the sum is taken over the arrival times of the spikes produced by the neurons j. The number of spikes arriving between t and $t + dt$ is $\nu_j(t)dt$. Therefore we have

$$V_i(t) = \sum_j \int_{-\infty}^{t} PSP_{ij}(t-s)\nu_j(s)\, ds = \sum_j \int_{-\infty}^{t} PSP_{ij}(t-s)S_j(V_j(s))\, ds,$$

or, equivalently

$$\nu_i(t) = S_i\left(\sum_j \int_{t_0}^{t} PSP_{ij}(t-s)\nu_j(s)\, ds\right). \tag{1}$$

The PSP_{ij}s can depend on several variables in order to account for instance for adaptation or learning.

The voltage-based model

The assumption, made in (Hopfield, 1984), is that the post-synaptic potential has the same shape no matter which presynaptic population caused it, the sign and amplitude may vary though. This leads to the relation

$$PSP_{ij}(t) = J_{ij} g_i(t).$$

g_i represents the unweighted shape (called a g-shape) of the postsynaptic potentials and J_{ij} is the strength of the postsynaptic potentials elicited by neuron j on neuron i. Thus we have

$$V_i(t) = \int_{t_0}^{t} g_i(t-s) \left(\sum_j J_{ij} \nu_j(s) \right) ds.$$

So far we have only considered the synaptic inputs to the neurons. We also assume that neuron i receives an external current density $I_i(t)$ and some noise $n_i(t)$ so that

$$V_i(t) = \int_{t_0}^{t} g_i(t-s) \left(\sum_j J_{ij} \nu_j(s) + I_i(s) + n_i(s) \right) ds. \tag{2}$$

We assume that the external current and the g-shapes satisfy $I_i = I_{p(i)}$, $g_i = g_{p(i)}$, $S_i = S_{p(i)}$, i.e. they only depend upon the neuron population. The noise model is described later. Finally we assume that $g_i = g_\alpha$ (where $\alpha = p(i)$) is the Green function of a linear differential equation of order k, i.e. satisfies

$$\sum_{l=0}^{k} a_\alpha^l(t) \frac{d^l g_\alpha}{dt^l}(t) = \delta(t). \tag{3}$$

We assume that the functions $a_\alpha^l(t)$ are continuous for $l = 0, \cdots, k$ and $\alpha = 1, \cdots, P$. We also assume $a_\alpha^k(t) \geq c > 0$ for all $t \in \mathbb{R}$, $\alpha = 1, \cdots, P$.

Known examples of g-shapes, see section 2.2.3 below, are $g_\alpha(t) = Ke^{-t/\tau} Y(t)$ ($k = 1, a_1(t) = \frac{1}{K}, a_0(t) = \frac{1}{K\tau}$)or $g_\alpha(t) = Kte^{-t/\tau} Y(t)$ ($k = 2, a_2(t) = \frac{1}{K}, a_1(t) = \frac{2}{K\tau}, a_0(t) = \frac{1}{K\tau}$), where Y is the Heaviside function.

We note D_α^k the corresponding differential operator, $D_\alpha^k g_\alpha = \delta$, and $\mathbf{D}_N^k$ the N-dimensional differential operator containing N_α copies of D_α^k, $\alpha = 1, \cdots, P$. We write (2) in vector form

$$\mathbf{V}^{(N)} = \mathbf{J}^{(N)} \mathrm{diag}(g_\alpha) * S^{(N)}(\mathbf{V}^{(N)}) + \mathrm{diag}(g_\alpha) * \mathbf{I}^{(N)} + \mathrm{diag}(g_\alpha) * \mathbf{n}^{(N)},$$

where $\mathrm{diag}(g_\alpha)$ is the N-dimensional diagonal matrix containing N_α copies of g_α, $\alpha = 1, \cdots, P$ and $*$ indicates the convolution operator. $S^{(N)}$ is the mapping $\mathbb{R}^N \to \mathbb{R}^N$ such that $S^{(N)}(\mathbf{V}^{(N)})_i = S_{p(i)}(\mathbf{V}_i^{(N)})$. We apply the operator $\mathbf{D}_N^k$ to both sides to obtain

$$\mathbf{D}_N^k \mathbf{V}^{(N)} = \mathbf{J}^{(N)} \cdot S^{(N)}(\mathbf{V}^{(N)}) + \mathbf{I}_V^{(N)} + \mathbf{n}_V^{(N)}, \tag{4}$$

which is a stochastic differential equation

$$d\left(\frac{d^{k-1}\mathbf{V}^{(N)}}{dt^{k-1}}\right) = \left(-\mathbf{D}_N^{k-1}\mathbf{V}^{(N)} + \mathbf{J}^{(N)} \cdot S^{(N)}(\mathbf{V}^{(N)}) + \mathbf{I}^{(N)}\right) dt + d\mathbf{n}_t^{(N)},$$

where $\mathbf{D}_N^{k-1}$ is obtained from the P differential operators of order $k-1$ $D_\alpha^{k-1} = \sum_{l=0}^{k-2} \frac{a_\alpha^l(t)}{a_\alpha^k(t)} \frac{d^l}{dt^l}$.

The activity-based model If we make the assumption that the shape of a PSP depends only on the nature of the presynaptic cell, that is

$$PSP_{ij} = J_{ij} g_j,$$

and define the activity as

$$A_j(t) = \int_{-\infty}^{t} g_j(t-s)\nu_j(s)\, ds,$$

multiplying both sides of equation (1) by $g_i(t-s)$ and integrating with respect to s, we obtain

$$A_i(t) = \int_{-\infty}^{t} g_i(t-s) S_i\left(\sum_j J_{ij} A_j(s) + I_i(s) + n_i(s)\right) ds,$$

where we have added an external current and a noise. If $p(i) = \alpha$, this yields

$$D_\alpha^k A_i = S_i\left(\sum_j J_{ij} A_j(t) + I_i(t) + n_i(t)\right),$$

and in terms of the N-dimensional vector $\mathbf{A}^{(N)}$

$$\mathbf{D}_N^k \mathbf{A}^{(N)} = S^{(N)}(\mathbf{J}^{(N)}\mathbf{A}^{(N)} + \mathbf{I}_A^{(N)} + \mathbf{n}_A^{(N)}). \tag{5}$$

Equivalence of the two models As a matter of fact these two equations are equivalent provided that $\mathbf{J}^{(N)}$ is invertible[1]. Indeed, let us use the change of variable $\mathbf{V}^{(N)} = \mathbf{J}^{(N)}\mathbf{A}^{(N)} + \mathbf{I}_A^{(N)} + \mathbf{n}_A^{(N)}$. We have, because $\mathbf{J}^{(N)}$ is not a function of time,

$$\mathbf{D}_N^k \mathbf{V}^{(N)} = \mathbf{J}^{(N)}\mathbf{D}_N^k \mathbf{A}^{(N)} + \mathbf{D}_N^k \mathbf{I}_A^{(N)} + \mathbf{D}_N^k \mathbf{n}_A^{(N)}.$$

Replacing $\mathbf{D}_N^k \mathbf{V}^{(N)}$ by this value in (4) we obtain

$$\mathbf{J}^{(N)}\mathbf{D}_N^k \mathbf{A}^{(N)} + \mathbf{D}_N^k \mathbf{I}_A^{(N)} + \mathbf{D}_N^k \mathbf{n}_A^{(N)} = \mathbf{J}^{(N)} \cdot S^{(N)}(\mathbf{J}^{(N)}\mathbf{A}^{(N)} + \mathbf{I}_A^{(N)} + \mathbf{n}_A^{(N)}) + \mathbf{I}_V^{(N)} + \mathbf{n}_V^{(N)}.$$

[1]Note that in the cases we treat in this paper, the matrix $\mathbf{J}^{(N)}$ is always almost surely invertible since it has non-degenerate Gaussian coefficients, and hence the equivalence in law will always be valid

Assuming that the matrix $\mathbf{J}^{(N)}$ is invertible yields

$$\mathbf{D}_N^k \mathbf{A}^{(N)} = S^{(N)}(\mathbf{J}^{(N)}\mathbf{A}^{(N)} + \mathbf{I}_A^{(N)} + \mathbf{n}_A^{(N)}) + (\mathbf{J}^{(N)})^{-1}\left(\mathbf{I}_V^{(N)} - \mathbf{D}_N^k \mathbf{I}_A^{(N)} + \mathbf{n}_V^{(N)} - \mathbf{D}_N^k \mathbf{n}_A^{(N)}\right).$$

Given the current $\mathbf{I}_V^{(N)}$ (respectively the noise $\mathbf{n}_V^{(N)}$), we can choose the current $\mathbf{I}_A^{(N)}$ (respectively the noise $\mathbf{n}_A^{(N)}$) solution of the linear differential equation $\mathbf{D}_N^k \mathbf{I}_A^{(N)} = \mathbf{I}_V^{(N)}$ (respectively $\mathbf{D}_N^k \mathbf{n}_A^{(N)} = \mathbf{n}_V^{(N)}$). Using the Green functions g_α, $\alpha = 1, \cdots, P$ this is equivalent to $\mathbf{I}_A^{(N)} = \mathrm{diag}(g_\alpha) * \mathbf{I}_V^{(N)}$ (respectively $\mathbf{n}_A^{(N)} = \mathrm{diag}(g_\alpha) * \mathbf{n}_V^{(N)}$).

The dynamics We introduce the $k-1$ N-dimensional vectors $\mathbf{V}^{(l)}(t) = [V_1^{(l)}, \cdots, V_N^{(l)}]^T$, $l = 1, \cdots, k-1$ of the lth-order derivative of $\mathbf{V}^{(N)}(t)$, and the Nk-dimensional vector

$$\widetilde{\mathbf{V}}^{(N)}(t) = \begin{bmatrix} \mathbf{V}^{(N)}(t) \\ \mathbf{V}^{(N)\,(1)}(t) \\ \vdots \\ \mathbf{V}^{(N)\,(k-1)}(t) \end{bmatrix}.$$

The N-neurons network is described by the Nk-dimensional vector $\widetilde{\mathbf{V}}^{(N)}(t)$. We consider the direct sum $\mathbb{R}^{Nk} = E^{(0)} \oplus \cdots \oplus E^{(k-1)}$, where each $E^{(l)} = \mathbb{R}^N$, $l = 0, \cdots, k-1$ and introduce the following notation: if $\mathbf{x}$ is a vector of $\mathbb{R}^{Nk}$, $\mathbf{x}_l$ is its component in $E^{(l)}$, $l = 0, \cdots, k-1$, an N-dimensional vector. In particular we have $\widetilde{\mathbf{V}}_l^{(N)}(t) = \mathbf{V}^{(N)\,(l)}(t)$ for $l = 0, \cdots, k-1$ with the convention that $\mathbf{V}^{(N)\,(0)} = \mathbf{V}^{(N)}$.

We now write the equations governing the time variation of the first $k-1$ vectors of $\widetilde{\mathbf{V}}^{(N)}(t)$, i.e. the derivatives of order 0, $\ldots k-2$ of $\mathbf{V}^{(N)}(t)$. These equation in effect determine the noise model. We write

$$d\widetilde{\mathbf{V}}_l^{(N)}(t) = \widetilde{\mathbf{V}}_{l+1}^{(N)}(t)\, dt + \mathbf{\Lambda}_l^{(N)} \cdot d\mathbf{W}_t^{(N)} \quad l = 0, \cdots, k-2, \tag{6}$$

where $\mathbf{\Lambda}_l^{(N)}$ is the $N \times N$ diagonal matrix $\mathrm{diag}(s_\alpha^l)$, where s_α^l, $\alpha = 1, \cdots, P$ is repeated N_α times, and $\mathbf{W}_t^{(N)}$ an N-dimensional standard Brownian process.

The equation governing the $(k-1)$th differential of the membrane potential has a linear part determined by the differential operator $\mathbf{D}^{k-1}$ and must account for the external inputs (deterministic and stochastic) and the activity of the neighbors, see (4). Keeping the same notations as before for the inputs and denoting by $\boldsymbol{\mathcal{L}}^{(N)}$ the $N \times Nk$ matrix describing the action of the neurons membrane potentials and their derivatives on the $(k-1)$th derivative of $\mathbf{V}$, we have:

$$d\widetilde{\mathbf{V}}_{k-1}^{(N)}(t) = \left(\boldsymbol{\mathcal{L}}^{(N)}(t) \cdot \widetilde{\mathbf{V}}^{(N)}(t) + \left(\mathbf{J}^{(N)} \cdot S^{(N)}(\widetilde{\mathbf{V}}_0^{(N)}(t))\right) + \mathbf{I}^{(N)}(t)\right) dt + \mathbf{\Lambda}_{k-1}^{(N)}(t) \cdot d\mathbf{W}_t^{(N)}, \tag{7}$$

where

$$\boldsymbol{\mathcal{L}}^{(N)} = \left[\begin{array}{ccc} \mathrm{diag}(a_\alpha^0(t)) & \cdots & \mathrm{diag}(a_\alpha^{k-1}(t)) \end{array} \right].$$

We define

$$\mathbf{L}^{(N)}(t) = \left[\begin{array}{cccc} 0 & \mathrm{Id}_N & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & \vdots & & \mathrm{Id}_N \\ \mathrm{diag}(a_\alpha^0(t)) & \mathrm{diag}(a_\alpha^1(t)) & \cdots & \mathrm{diag}(a_\alpha^{k-1}(t)) \end{array} \right],$$

where Id_N is the $N \times N$ identity matrix. We also denote by:

$$\widetilde{\mathbf{U}}_t^{(N)} = \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ \mathbf{J}^{(N)} \cdot S^{(N)}(\widetilde{\mathbf{V}}_0(t)) \end{array} \right] \text{ and } \widetilde{\mathbf{I}}_t^{(N)} = \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ \mathbf{I}^{(N)}(t) \end{array} \right].$$

The full equation satisfied by $\widetilde{\mathbf{V}}^{(N)}$ can be written:

$$d\widetilde{\mathbf{V}}^{(N)}(t) = \left(\mathbf{L}^{(N)}(t)\widetilde{\mathbf{V}}^{(N)}(t) + \widetilde{\mathbf{U}}_t^{(N)} + \widetilde{\mathbf{I}}_t^{(N)} \right) dt + \boldsymbol{\Lambda}^{(N)}(t) \cdot d\mathbf{W}_t^{(N)}, \tag{8}$$

where the $kN \times kN$ matrix $\boldsymbol{\Lambda}^{(N)}(t)$ is equal to $\mathrm{diag}(\boldsymbol{\Lambda}_0^{(N)}, \cdots, \boldsymbol{\Lambda}_{k-1}^{(N)})$.

Note that the kth-order differential equation describing the time variation of the membrane potential of each neuron contains a noise term which is a linear combination of various integrated Brownian processes (up to the order $k-1$) as shown in the following formula which is derived from (6) and (7).

$$\begin{aligned} d\widetilde{\mathbf{V}}_{k-1}^{(N)}(t) = & \left(\sum_{l=0}^{k-1} \mathbf{L}_l^{(N)}(t)\widetilde{\mathbf{V}}_l^{(N)} \right) dt + \\ & \left(\sum_{l=0}^{k-2} \mathbf{L}_l^{(N)}(t) \left(\sum_{h=0}^{k-l-2} \int_0^t \int_0^{s_1} \cdots \int_0^{s_{h-1}} \boldsymbol{\Lambda}_{l+h}^{(N)}(s_h) d\mathbf{W}_{s_h} d_{s_{h-1}} \cdots ds_0 \right) \right) dt + \\ & \boldsymbol{\Lambda}_{k-1}^{(N)}(t) d\mathbf{W}_t. \end{aligned}$$

Comparing with equation (4) we see that the noise $\mathbf{n}^{(N)}\, dt$ is a weighted sum of Brownian and integrated Brownian processes.

2.2 The Mean Field equations

2.2.1 General derivation of the mean field equation

The connectivity weight J_{ij} are modeled as independent Gaussian random variables. Their distribution depends only on the population pair $\alpha = p(i), \beta = p(j)$, and on the total number of neurons N_β of population β:

$$J_{ij} \sim \mathcal{N}\Big(\frac{\bar{J}_{\alpha\beta}}{N_\beta}, \frac{\sigma_{\alpha\beta}}{\sqrt{N_\beta}}\Big).$$

We are interested in the limit law when $N \to \infty$ of the vector $\mathbf{V}^{(N)}$ under the joint law of the connectivities and the Brownian motions, which we call the mean field limit. This law can be described by a set of P equations, the mean field equations. As mentioned in the introduction these equations can be derived in several ways, either heuristically as in the work of Amari (Amari, 1972; Amari et al, 1977), Sompolinsky (Crisanti et al, 1990; Sompolinsky et al, 1988), and Cessac (Cessac, 1995; Samuelides and Cessac, 2007), or rigorously as in the work of Benarous and Guionnet (Ben-Arous and Guionnet, 1995, 1997; Guionnet, 1997). We derive them here in a pedestrian way, prove that they are well-posed, and provide an algorithm for computing their solution.

The effective description of the network population by population is possible because the neurons in each population are interchangeable, i.e. have the same probability distribution under the joint law of the multidimensional Brownian motion and the connectivity weights. This is the case because of the form of equation (8).

The Mean Field equations We note $C([t_0,T],\mathbb{R}^P)$ (respectively $C((-\infty,T],\mathbb{R}^P)$) the set of continuous functions from the real interval $[t_0,T]$ (respectively $(-\infty,T]$) to $\mathbb{R}^P$. By assigning a probability to subsets of such functions, a continuous stochastic process X defines a positive measure of unit mass on $C([t_0,T],\mathbb{R}^P)$ (respectively $C((-\infty,T],\mathbb{R}^P)$). This set of positive measures of unit mass is noted $\mathcal{M}_1^+(C([t_0,T],\mathbb{R}^P))$ (respectively $\mathcal{M}_1^+(C((-\infty,T],\mathbb{R}^P))$.

We now define a process of particular importance for describing the limit process: the effective interaction process.

Definition 2.1 (Effective Interaction Process)**.** Let $X \in \mathcal{M}_1^+(C([t_0,T],\mathbb{R}^P))$ (resp. $\mathcal{M}_1^+(C((-\infty,T],\mathbb{R}^P))$ be a given stochastic process. The effective interaction term is the Gaussian process $\mathbf{U}^X \in \mathcal{M}_1^+(C([t_0,T],\mathbb{R}^{P\times P}))$, (resp. $\mathcal{M}_1^+(C((-\infty,T],\mathbb{R}^{P\times P}))$ statistically independent of the external noise $(\mathbf{W}_t)_{t\geq t_0}$ and of the initial condition X_{t_0} (when $t_0 > -\infty$), defined by:

$$\begin{cases} \mathbb{E}\left[U^X_{\alpha\beta}(t)\right] = \bar{J}_{\alpha\beta} m^X_{\alpha\beta}(t) \ \text{ where } \ m^X_{\alpha\beta}(t) \stackrel{\text{def}}{=} \mathbb{E}[S_{\alpha\beta}(X_\beta(t))]; \\ \mathrm{Cov}(U^X_{\alpha\beta}(t), U^X_{\alpha\beta}(s)) = \sigma^2_{\alpha\beta}\Delta^X_{\alpha\beta}(t,s) \text{ where} \\ \qquad\qquad \Delta^X_{\alpha\beta}(t,s) \stackrel{\text{def}}{=} \mathbb{E}\Big[S_{\alpha\beta}(X_\beta(t))S_{\alpha\beta}(X_\beta(s))\Big]; \\ \mathrm{Cov}(U^X_{\alpha\beta}(t), U^X_{\gamma\delta}(s)) = 0 \text{ if } \alpha\neq\gamma \text{ or } \beta\neq\delta. \end{cases} \tag{9}$$

Choose P neurons $i_1,\ldots,i_P$, one in each population (neuron i_α belongs to the population α). Then it can be shown, using either a heuristic argument or large deviations techniques (see appendix A), that the sequence of processes $\left(\widetilde{\mathbf{V}}^{(N)}(t) = [\tilde{V}^{(N)}_{i_1}(t),\ldots,\tilde{V}^{(N)}_{i_P}(t)]^T_{t\geq t_0}\right)_{N\geq 1}$ converges in law to the process $\widetilde{\mathbf{V}}(t) = [\tilde{V}_1(t),\ldots,\tilde{V}_P(t)]^T_{t\geq t_0}$ solution of the following mean field equation:

$$d\widetilde{\mathbf{V}}(t) = \Big(\mathbf{L}(t)\widetilde{\mathbf{V}}(t) + \widetilde{\mathbf{U}}^{\mathbf{V}}_t + \widetilde{\mathbf{I}}(t)\Big)\,dt + \mathbf{\Lambda}(t)\cdot d\mathbf{W}_t, \tag{10}$$

where $\widetilde{\mathbf{V}}$ is a kP-dimensional vector containing the P-dimensional vector $\mathbf{V}$ and its $k-1$ derivatives, and $\mathbf{L}$ is the $Pk \times Pk$ matrix

$$\mathbf{L}(t) = \begin{bmatrix} 0_{P\times P} & \mathrm{Id}_P & \cdots & 0_{P\times P} \\ 0_{P\times P} & 0_{P\times P} & \ddots & 0_{P\times P} \\ \vdots & \vdots & & \mathrm{Id}_P \\ \mathbf{L}_0(t) & \mathbf{L}_1(t) & \cdots & \mathbf{L}_{k-1}(t) \end{bmatrix},$$

where Id_P is the $P \times P$ identity matrix and $0_{P\times P}$ the null $P \times P$ matrix. $(\mathbf{W}_t)$ is a kP-dimensional standard Brownian process and:

$$\widetilde{\mathbf{U}}_t^{\mathbf{V}} = \begin{bmatrix} 0_P \\ \vdots \\ 0_P \\ \mathbf{U}_t^{\mathbf{V}} \cdot \mathbf{1} \end{bmatrix} \quad \widetilde{\mathbf{I}}(t) = \begin{bmatrix} 0_P \\ \vdots \\ 0_P \\ \mathbf{I}(t) \end{bmatrix} \quad \mathbf{\Lambda}(t) = \mathrm{diag}(\mathbf{\Lambda}_0(t), \cdots, \mathbf{\Lambda}_{k-1}(t)).$$

The matrices $\mathbf{L}_0, \cdots, \mathbf{L}_{k-1}$ (respectively $\mathbf{\Lambda}_0, \cdots, \mathbf{\Lambda}_{k-1}$) are obtained by selecting the same P rows and P columns of the matrices $\mathbf{L}_0^{(N)}, \cdots, \mathbf{L}_{k-1}^{(N)}$ (respectively $\mathbf{\Lambda}_0^{(N)}, \cdots, \mathbf{\Lambda}_{k-1}^{(N)}$) corresponding to P neurons in different populations, $(\mathbf{U}_t^{\mathbf{V}})$ is the effective interaction process associated with $\mathbf{V}$, and $\mathbf{I}(\cdot)$ is the P-dimensional external current.

To proceed further we formally integrate the equation using the flow, or resolvent, of the equation, noted $\Phi_L(t, t_0)$ (see appendix B), and we obtain, since we assumed $\mathbf{L}$ continuous, an implicit representation of $\mathbf{V}$:

$$\widetilde{\mathbf{V}}(t) = \Phi_L(t, t_0)\widetilde{\mathbf{V}}(t_0) + \int_{t_0}^{t} \Phi_L(t, s) \cdot \left(\widetilde{\mathbf{U}}_s^{\mathbf{V}} + \widetilde{\mathbf{I}}(s)\right) ds + \int_{t_0}^{t} \Phi_L(t, s) \cdot \mathbf{\Lambda}(s) d\mathbf{W}_s \quad (11)$$

We now introduce for future reference a simpler model which is quite frequently used in the description on neural networks.

2.2.2 Example: The Simple Model

In the Simple Model, each neuron membrane potential decreases exponentially to its rest value if it receives no input, with a time constant τ_α depending only on the population. The noise is modeled by an independent Brownian process per neuron whose standard deviation is the same for all neurons belonging to a given population.

Hence the dynamics of a given neuron i from population α of the network reads:

$$dV_i^{(N)}(t) = \Big[-\frac{V_i^{(N)}(t)}{\tau_\alpha} + \sum_{\beta=1}^{P}\sum_{j=1}^{N_\beta} J_{ij} S_{\alpha\beta}\big(V_j^{(N)}(t)\big) + I_\alpha(t)\Big]\, dt + s_\alpha dW_{(i)}(t). \quad (12)$$

This is a special case of equation (10) where $k = 1$ and $\mathbf{L} = -\mathrm{diag}(\frac{1}{\tau_1}, \cdots, \frac{1}{\tau_P})$, $\Phi_L(t, t_0) = \mathrm{diag}(e^{-(t-t_0)/\tau_1}, \cdots, e^{-(t-t_0)/\tau_P})$, and $\mathbf{\Lambda} = \mathrm{diag}(s_1, \cdots, s_P)$. The corresponding mean field equation reads:

$$dV_\alpha(t) = \Big(-\frac{V_\alpha}{\tau_\alpha}(t) + \sum_{\beta=1}^{P} U_{\alpha\beta}^V(t) + I_\alpha(t)\Big)\, dt + s_\alpha dW_\alpha(t), \quad \forall \alpha \in \{1, \ldots, P\}, \quad (13)$$

where the processes $(W_\alpha(t))_{t\geq 0}$ are independent standard Brownian motions, $\mathbf{U}^V(t) = (U_{\alpha\beta}^V(t);\ \alpha, \beta \in \{1, \ldots, P\})_t$ is the effective interaction term.

This equation can be integrated implicitly and we obtain the following integral representation of the process $V_\alpha(t)$:

$$V_\alpha(t) = e^{-(t-t_0)/\tau_\alpha} V_\alpha(t_0) + \int_{t_0}^{t} e^{-(t-s)/\tau_\alpha} \Big(\sum_{\beta=1}^{P} U_{\alpha\beta}^V(s) + I_\alpha(s)\Big)\, ds + s_\alpha \int_{t_0}^{t} e^{-(t-s)/\tau_\alpha} dW_\alpha(s)\Big) \quad (14)$$

where t_0 is the initial time. It is an implicit equation on the probability distribution of $\mathbf{V}(t)$, a special case of (11).

2.2.3 The Jansen and Rit's model

One of the motivations of this study is to characterize the global behavior of an assembly of neurons in particular to get a better understanding of non-invasive cortical signals like EEG or MEG. One of the classical models of neural masses is Jansen and Rit's mass model (Jansen and Rit, 1995), in short the JR model (see figure 1).

The model features a population of pyramidal neurons (central part of figure 1.a.) that receives excitatory and inhibitory feedback from local inter-neurons and an excitatory input from neighboring cortical units and sub-cortical structures such as the thalamus. The excitatory input is represented by an arbitrary average firing rate $p(t)$ that can be stochastic (accounting for a non specific background activity) or deterministic, accounting for some specific activity in other cortical units. The transfer functions h_e and h_i of figure 1 convert the average firing rate describing the input to a population

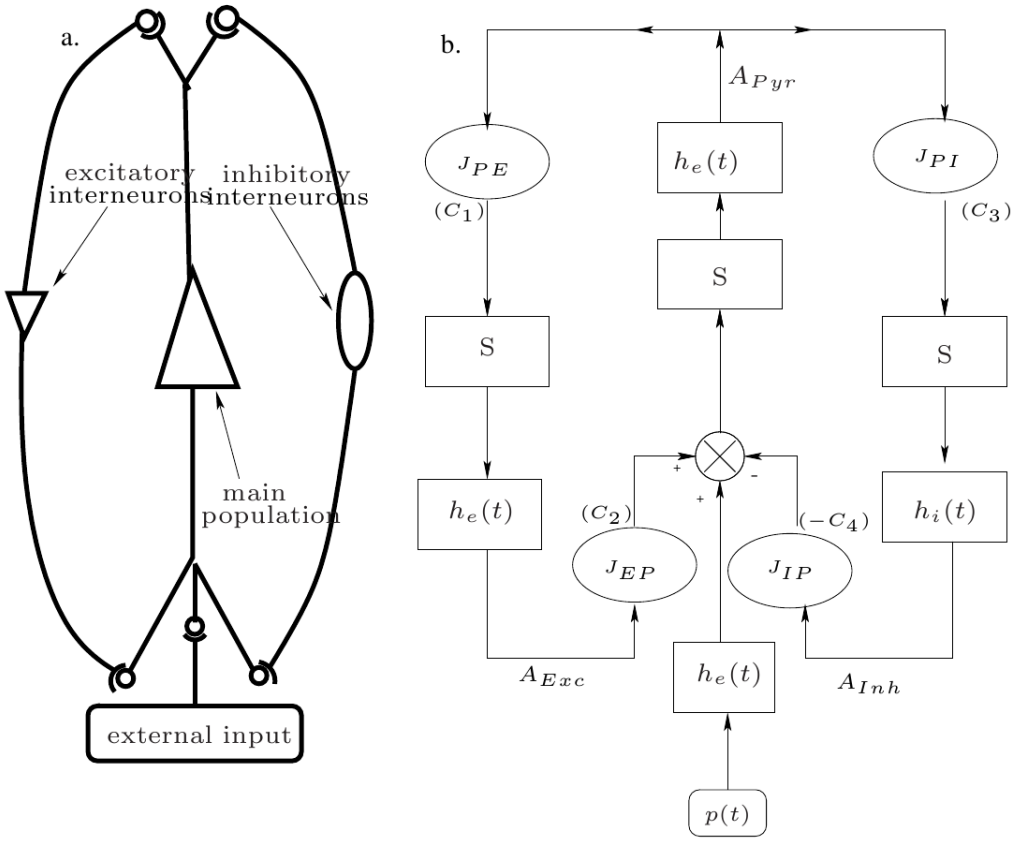


Figure 1: a. Neural mass model: a population of pyramidal cells interacts with two populations of inter-neurons: an excitatory one and an inhibitory one. b. Block representation of the model. The *h* boxes account for the synaptic integration between neuronal populations. S boxes simulate cell bodies of neurons by transforming the membrane potential of a population into an output firing rate. The coefficient $J_{\alpha\beta}$ is the random synaptic efficiency of population β on population α (P is the pyramidal population, E the excitatory and I the inhibitory ones), and the constants (C_i) model the mean strength of the synaptic connections between populations (it is the mean of the Js).

into an average excitatory or inhibitory post-synaptic potential (EPSP or IPSP). They correspond to the synaptic integration.

In the model introduced originally by Jansen and Rit, the connectivity weights were assumed to be constant, equal to their mean value (it is the constants C_i, $i = 1 \dots 4$ in figure 1). Nevertheless, there exists a variability on these coefficients, and as we will see in the sequel, the effect of the connectivity variability impacts the solution at the level of the neural mass. Statistical properties

of the connectivities have been studied in details for instance in (Braitenberg and Schüz, 1998). In our model we consider these connectivities as independent Gaussian random variables of mean and standard deviation equal to the ones found in (Braitenberg and Schüz, 1998).

We now use diagram 1 to derive the membrane potential expressions. We consider a network of N neurons belonging to the three populations described. We denote by P (resp E, I) the pyramidal (respectively excitatory, inhibitory) populations. We choose in population P (respectively populations E, I) a particular pyramidal neuron (respectively excitatory, inhibitory interneuron) indexed by i_{pyr} (respectively i_{exc}, i_{inh}). The equations of their activity variable read:

$$\begin{cases} A^N_{i_{pyr}} &= h_e * S(\sum_{j\ Exc} J_{ij} A^N_j + h_e * p(\cdot) + \sum_{j\ Inh} J_{ij} A^N_j) \\ A^N_{i_{exc}} &= h_e * S(\sum_{j\ Pyr} J_{ij} A^N_j) \\ A^N_{i_{inh}} &= h_i * S(\sum_{j\ Pyr} J_{ij} A^N_j) \end{cases}$$

This is therefore an activity-based model. As stated before, it is equivalent via a change of variable to a voltage-based model, with the same connectivity matrix, the same intrinsic dynamics, and modified inputs (see section 2.1.1).

In the mean field limit, denoting by A_P (respectively A_E, A_I) the activity of the pyramidal neurons (resp excitatory, inhibitory interneurons), we obtain the following activity equations:

$$\begin{cases} A_P &= h_e * S(U_{PE} + h_e * p + U_{PI}) \\ A_E &= h_e * S(U_{EP}) \\ A_I &= h_e * S(U_{IP}) \end{cases} \tag{15}$$

where $\mathbf{U} = (U_{ij})_{i,j\in\{P,E,I\}}$ is the effective interaction process associated with this problem, i.e. a Gaussian process of means:

$$\begin{cases} \mathbb{E}[U_{EP}] &= \bar{J}_{EP}\mathbb{E}[A_E] \\ \mathbb{E}[U_{IP}] &= \bar{J}_{IP}\mathbb{E}[A_I] \\ \mathbb{E}[U_{PI}] &= \bar{J}_{PI}\mathbb{E}[A_P] \\ \mathbb{E}[U_{PE}] &= \bar{J}_{PE}\mathbb{E}[A_P] \end{cases}$$

and whose covariance matrix can be deduced from (9). The voltage-based model can be deduced from this activity-based description using a simple change of variable as stated previously. Note that the change of variable is possible since the activity current I_A is equal to $h_e * p$ and, as shown in section 2.1.1, I_A is smooth enough so that we can apply to it the suitable differential operator. p is the corresponding voltage current I_V.

Let us now instantiate the synaptic dynamics and compare the mean field equation with Jansen's population equations (sometimes improperly called also mean field equations).

The simplest model of synaptic integration is a first-order integration, which yields exponential post-synaptic potentials:

$$h(t) = \begin{cases} \alpha e^{-\beta t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

that satisfies the following differential equations

$$\dot{h}(t) = -\beta h(t) + \alpha\delta(t),$$

In these equations β is the time constant of the synaptic integration and α the synaptic efficiency. The coefficients named α and β are the same for the pyramidal and the excitatory population, and different from the ones of the inhibitory synapse. In the pyramidal or excitatory (respectively the inhibitory) case we have $\alpha = A$, $\beta = a$ (respectively $\alpha = B$, $\beta = b$). Eventually, the sigmoid functions are the same whatever the populations, and is given by

$$S(v) = \frac{\nu_{\max}}{1 + e^{r(v_0 - v)}},$$

$\nu_{\max}$ is the maximum firing rate, and v_0 is a voltage reference.

With this synaptic dynamics we obtain the first-order Jansen and Rit's equation:

$$\begin{cases} \frac{dA_P}{dt}(t) &= -a\, A_P(t) + A\, S(U_{PE} + U_{P\,I} + h_e * p(t)) \\ \frac{dA_E}{dt}(t) &= -a\, A_E(t) + A\, S(U_{EP}) \\ \frac{dA_I}{dt}(t) &= -b\, A_I(t) + B\, S(U_{IP}) \end{cases} . \tag{16}$$

while the "original" Jansen and Rit's equation (Grimbert and Faugeras, 2006; Jansen and Rit, 1995) reads:

$$\begin{cases} \frac{dA_P}{dt}(t) &= -a\, A_P(t) + A\, S(C_2\, A_E(t) - C_4\, A_I(t) + h_e * p(t)) \\ \frac{dA_E}{dt}(t) &= -a\, A_E(t) + A\, S(C_1\, A_P(t)) \\ \frac{dA_I}{dt}(t) &= -b\, A_I(t) + B\, S(C_3\, A_P(t)) \end{cases} . \tag{17}$$

Hence the original JR equation amounts to computing the expectation of the activity in each population and to assume that

$$\mathbb{E}\left[S(U_{P\,E} + U_{P\,I} + h_e * p(t))\right] = S(\mathbb{E}\left[U_{P\,E} + U_{P\,I} + h_e * p\right]),$$

which is a quite sharp assumption given that the sigmoidal function is nonlinear.

A higher order model was introduced to better account for the synaptic integration and to better reproduce the characteristics of real EPSPs and IPSPs by van Rotterdam and colleagues (van Rotterdam et al, 1982). In this model the PSP satisfies a second order differential equation:

$$h(t) = \begin{cases} \alpha\beta t e^{-\beta t} & t \geq 0 \\ 0 & t < 0 \end{cases} ,$$

solution of the differential equation $\ddot{y}(t) = \alpha\beta\delta(t) - 2\beta\dot{y}(t) - \beta^2 y(t)$. With this type of synaptic integration, we obtain the following mean field equations:

$$\begin{cases} \frac{d^2A_P}{dt^2}(t) &= AaS(U_{PE} + U_{PI} + h_e * p(t)) - 2a\frac{dA_P}{dt}(t) - a^2A_P(t) \\ \frac{d^2A_E}{dt^2}(t) &= AaS(U_{EP}) - 2a\frac{dA_E}{dt}(t) - a^2A_E(t) \\ \frac{d^2A_I}{dt^2}(t) &= BbS(U_{IP}) - 2b\frac{dA_I}{dt}(t) - b^2A_I(t) \end{cases} \tag{18}$$

while the original system satisfies the equations:

$$\begin{cases} \frac{d^2A_P}{dt^2}(t) &= AaS(C_2A_E(t) - C_4A_I(t) + h_e * p(t)) - 2a\frac{dA_P}{dt}(t) - a^2A_P(t) \\ \frac{d^2A_E}{dt^2}(t) &= AaS(C_1A_P(t)) - 2a\frac{dA_E}{dt}(t) - a^2A_E(t) \\ \frac{d^2A_I}{dt^2}(t) &= BbC_4S(C_3A_P(t)) - 2b\frac{d^2A_I}{dt^2}(t) - b^2A_I(t) \end{cases} \tag{19}$$

Here again, going from the mean field equations (18) to the neural mass model (19) consists in studying the equation of the mean of the process given by (18) and commuting the sigmoidal function with the expectation.

Note that the introduction of higher order synaptic integrations results in richer behaviors. For instance, Grimbert and Faugeras (Grimbert and Faugeras, 2006) showed that some bifurcations can appear in the second-order JR model giving rise to epileptic like oscillations and alpha activity, that do not appear in the first order model.

3 Existence and uniqueness of solutions in finite time

The mean field equation (11) is an implicit equation of the stochastic process $(V(t))_{t\geq t_0}$. We prove in this section that under some mild assumptions this implicit equation has a unique solution. This solution is a fixed point in the set $\mathcal{M}_1^+(C([t_0,T],\mathbb{R}^{kP}))$ of kP-dimensional processes. We construct a sequence of processes and prove that it converges in distribution toward this fixed point.

We denote by $\mathcal{X}$ the set of random variables (r.v.) with values in $\mathbb{R}^{kP}$. We first recall some results on the convergence of random variables and stochastic processes.

3.1 Convergence of Gaussian processes

We recall the following result from (Bogachev, 1998).

Theorem 3.1. *Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of kP-dimensional Gaussian processes defined on $[t_0,T]$ or on an unbounded interval of $\mathbb{R}^2$. The sequence converges to a Gaussian process X if and only if the following three conditions are satisfied:*

- *The sequence $\{X_n\}_{n=1}^{\infty}$ is uniformly tight.*
- *The sequence $\mu_n(t)$ of the mean functions converges for the uniform norm.*

[2] In (Bogachev, 1998, Chapter 3.8), the property is stated whenever the mean and covariance are defined on a separable Hilbert space.

- *The sequence C_n of the covariance operators converges for the uniform norm.*

We now define such a sequence of Gaussian processes.

Let us fix Z_0, a kP-dimensional Gaussian random variable, independent of the Brownian.

Definition 3.1. Let X an element of $\mathcal{M}_1^+(C([t_0,T],\mathbb{R}^{kP}))$. Let F_k be the function such that

$$F_k(X)_t = \Phi_L(t,t_0)\cdot Z_0 + \int_{t_0}^t \Phi_L(t,s)\cdot\left(\widetilde{\mathbf{U}}_s^X + \widetilde{\mathbf{I}}(s)\right)ds + \int_{t_0}^t \Phi_L(t,s)\cdot\mathbf{\Lambda}(s)d\mathbf{W}_s$$

where $\widetilde{\mathbf{U}}_s^X$ and $\widetilde{\mathbf{I}}(s)$ are defined in section 2.

Note that, by definition, the random process $(F_k(X))_{t\in[t_0,T]}$, $k\geq 1$ is the sum of a deterministic function (defined by the external current) and three independent random processes defined by the initial condition, the interaction between neurons, and the external noise.
Let X be a given stochastic process of $\mathcal{M}_1^+(C([t_0,T],\mathbb{R}^{kP}))$ such that $X_{t_0}=Z_0$. We define the sequence of processes $\{X_n\}_{n=0}^\infty \in \mathcal{M}_1^+(C([t_0,T],\mathbb{R}^{kP})$ by:

$$\begin{cases} X_0 & = X \\ X_{n+1} & = F_k(X_n) = F_k^{(n)}(X_0). \end{cases} \tag{20}$$

In the remaining of this section we show that the sequence of processes $\{F_k^{(n)}(X)\}_{n=0}^\infty$ converges in distribution toward the unique fixed-point Y of F_k.

3.2 Existence and uniqueness of solution for the mean field equations

The following upper and lower bounds are used in the sequel.

Lemma 3.2. We consider the Gaussian process $((\mathbf{U}_t^X\cdot\mathbf{1})_t)_{t\in[t_0,T]}$. $\mathbf{U}^X$ is defined in 2.1 and $\mathbf{1}$ is the P-dimensional vector with all coordinates equal to 1. We have

$$\left\|\mathbb{E}\left[\mathbf{U}_t^X\cdot\mathbf{1}\right]\right\|_\infty \leq \mu \stackrel{\text{def}}{=} \max_\alpha \sum_\beta |\bar{J}_{\alpha\beta}|\,\|S_{\alpha\beta}\|_\infty \tag{21}$$

for all $t_0\leq t\leq T$. The maximum eigenvalue of its covariance matrix is upperbounded by $\sigma_{\max}^2 \stackrel{\text{def}}{=} \max_\alpha \sum_\beta \sigma_{\alpha\beta}^2\,\|S_{\alpha\beta}\|_\infty^2$ where $\|S_{\alpha\beta}\|_\infty$ is the supremum of the absolute value of $S_{\alpha\beta}$.

Proof. The proof is straightforward from definition 3.1. We also note $\sigma_{\min}^2 \stackrel{\text{def}}{=} \min_{\alpha,\beta}\sigma_{\alpha\beta}^2$. □

The proof of existence and uniqueness of solution, and of the convergence of the sequence (20) is in two main steps. We first prove that the sequence of Gaussian processes $\{F_k^{(n)}(X)\}_{n=0}^\infty$, $k\geq 1$ is uniformly tight by proving that Kolmogorov's criterion for tightness holds. This takes care of condition 1) in theorem 3.1. We next prove that the sequences of the mean functions and covariance operators are Cauchy sequences for the uniform norms, taking care of conditions 2) and 3).

Theorem 3.3. *The sequence of processes* $\left\{F_k^{(n)}(X)\right\}_{n=0}^{\infty}$, $k \geq 1$ *is uniformly tight.*

Proof. We use Kolmogorov's criterion for tightness and do the proof for $k = 1$, the case $k > 1$ is similar. If we assume that $n \geq 1$ and $s < t$ we have

$$
\begin{aligned}
F_1^{(n)}(X)_t - F_1^{(n)}(X)_s &= (\Phi_L(t,t_0) - \Phi_L(s,t_0))X_{t_0} \\
&+ \int_{t_0}^{s} (\Phi_L(t,s) - Id)\Phi_L(s,u)\mathbf{U}_u^{F_1^{(n-1)}(X)} \cdot \mathbf{1}\, du + \int_s^t \Phi_L(t,u)\mathbf{U}_u^{F_1^{(n-1)}(X)} \cdot \mathbf{1}\, du \\
&+ \int_{t_0}^{s} (\Phi_L(t,s) - Id)\Phi_L(s,u)\mathbf{\Lambda}(u)\, d\mathbf{W}_u + \int_s^t \Phi_L(t,u)\mathbf{\Lambda}(u)\, d\mathbf{W}_u \\
&+ \int_{t_0}^{s} (\Phi_L(t,s) - Id)\Phi_L(s,u)\mathbf{I}(u)\, du + \int_s^t \Phi_L(t,u)\mathbf{I}(u)\, du
\end{aligned}
$$

and therefore (Cauchy-Schwarz and Jensen's inequalities):

$$
\begin{aligned}
\frac{1}{7}\|F_1^{(n)}(X)_t - F_1^{(n)}(X)_s\|^2 &\leq \|\Phi_L(t,t_0) - \Phi_L(s,t_0)\|^2\|\mathbf{X}_{t_0}\|^2 \\
&+ (s-t_0)\|\Phi_L(t,s) - Id\|^2 \int_{t_0}^{s} \|\Phi_L(s,u)\|^2 \|\mathbf{U}_u^{F_1^{(n-1)}(X)} \cdot \mathbf{1}\|^2\, du \\
&+ (t-s)\int_s^t \|\Phi_L(t,u)\|^2\|\mathbf{U}_u^{F_1^{(n-1)}(X)} \cdot \mathbf{1}\|^2\, du \\
&+ \left\|\int_{t_0}^{s} \Phi_L(s,u)(\Phi_L(t,s) - Id)\mathbf{\Lambda}(u)\, d\mathbf{W}_u\right\|^2 + \left\|\int_s^t \Phi_L(t,u)\mathbf{\Lambda}(u)\, d\mathbf{W}_u\right\|^2 \\
&+ (s-t_0)^2\|\Phi_L(t,s) - Id\|^2 I_{\max}^2 \sup_{u\in[t_0,s]} \|\Phi_L(s,u)\|^2 \\
&+ (t-s)^2 I_{\max}^2 \sup_{u\in[s,t]} \|\Phi_L(t,u)\|^2.
\end{aligned}
$$

Because $\|\Phi_L(t,t_0) - \Phi_L(s,t_0)\| \leq |t-s|\|\mathbf{L}\|$ we see that all terms in the righthand side of the inequality but the two involving the Brownian motion are of the order of $(t-s)^2$. We raise again both sides to the second power, use the Cauchy-Schwarz inequality again, and take the expected value:

$$
\begin{aligned}
\frac{1}{7^3}\mathbb{E}\left[\|F_1^{(n)}(X)_t - F_1^{(n)}(X)_s\|^4\right] &\leq \|\Phi_L(t,t_0)-\Phi_L(s,t_0)\|^4\mathbb{E}\left[\|\mathbf{X}_{t_0}\|^4\right] \\
&+ (s-t_0)^3\|\Phi_L(t,s)-Id\|^4\int_{t_0}^{s}\|\Phi_L(s,u)\|^4\mathbb{E}\left[\|\mathbf{U}_u^{F_1^{(n-1)}(X)}\cdot\mathbf{1}\|^4\right]du \\
&+ (t-s)^3\int_{s}^{t}\|\Phi_L(t,u)\|^4\mathbb{E}\left[\|\mathbf{U}_u^{F_1^{(n-1)}(X)}\cdot\mathbf{1}\|^4\right]du \\
&+ \mathbb{E}\left[\left\|\int_{t_0}^{s}\Phi_L(s,u)(\Phi_L(t,s)-Id)\mathbf{\Lambda}(u)\,d\mathbf{W}_u\right\|^4\right] \\
&+ \mathbb{E}\left[\left\|\int_{s}^{t}\Phi_L(t,u)\mathbf{\Lambda}(u)\,d\mathbf{W}_u\right\|^4\right] \\
&+ (s-t_0)^4\|\Phi_L(t,s)-Id\|^4\sup_{u\in[t_0,s]}\|\Phi_L(s,u)\|^4 I_{\max}^4 \\
&+ (t-s)^4 I_{\max}^4\sup_{u\in[s,t]}\|\Phi_L(t,u)\|^4.
\end{aligned}
\tag{22}
$$

Remember that $\mathbf{U}_u^{F_1^{(n-1)}(X)}\cdot\mathbf{1}$ is a P-dimensional diagonal Gaussian process, noted $\mathbf{Y}_u$, therefore:

$$
\mathbb{E}\left[\|\mathbf{Y}_u\|^4\right] = \sum_{\alpha}\mathbb{E}\left[Y_\alpha(u)^4\right] + \sum_{\alpha_1\neq\alpha_2}\mathbb{E}\left[Y_{\alpha_1}^2(u)\right]\mathbb{E}\left[Y_{\alpha_2}^2(u)\right].
$$

The second order moments are upperbounded by some regular function of μ and $\sigma_{\max}$ (defined in lemma 3.2) and, because of the properties of Gaussian integrals, so are the fourth order moments.

Let us now evaluate $\mathbb{E}\left[\left\|\int_a^b \mathbf{A}(u)\,d\mathbf{W}_u\right\|^4\right]$ for some $P\times P$ matrix $\mathbf{A}$. We have

$$
\begin{aligned}
\mathbb{E}\left[\left\|\int_a^b \mathbf{A}(u)\,d\mathbf{W}_u\right\|^4\right] &= \mathbb{E}\left[\left(\left\|\int_a^b \mathbf{A}(u)\,d\mathbf{W}_u\right\|^2\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^{P}\left(\sum_{j=1}^{P}\int_a^b A_{ij}(u)\,dW_u^j\right)\left(\sum_{k=1}^{P}\int_a^b A_{ik}(u)\,dW_u^k\right)\right)^2\right] \\
&= \sum_{i_1,i_2,j_1,j_2,k_1,k_2}\mathbb{E}\left[\int_a^b A_{i_1j_1}(u)\,dW_u^{j_1}\int_a^b A_{i_1k_1}(u)\,dW_u^{k_1}\int_a^b A_{i_2j_2}(u)\,dW_u^{j_2}\int_a^b A_{i_2k_2}(u)\,dW_u^{k_2}\right].
\end{aligned}
$$

Because of the properties of the Brownian process, the last term is the sum of three types of terms:

$$\sum_{i_1,i_2} \mathbb{E}\left[\left(\int_a^b A_{i_1 j}(u)\, dW_u^j\right)^2 \left(\int_a^b A_{i_2 j}(u)\, dW_u^j\right)^2\right] \leq \sum_{i_1,i_2} \mathbb{E}\left[\left(\int_a^b A_{i_1 j}(u)\, dW_u^j\right)^4\right]^{1/2} \mathbb{E}\left[\left(\int_a^b A_{i_2 j}(u)\, dW_u^j\right)^4\right]^{1/2},$$

and

$$\sum_{i_1,i_2,j_1\neq j_2} \mathbb{E}\left[\left(\int_a^b A_{i_1 j_1}(u)\, dW_u^{j_1}\right)^2 \left(\int_a^b A_{i_2 j_2}(u)\, dW_u^{j_2}\right)^2\right] = \sum_{i_1,i_2,j_1\neq j_2} \mathbb{E}\left[\left(\int_a^b A_{i_1 j_1}(u)\, dW_u^{j_1}\right)^2\right] \mathbb{E}\left[\left(\int_a^b A_{i_2 j_2}(u)\, dW_u^{j_2}\right)^2\right],$$

and

$$\sum_{i_1,i_2,j_1\neq j_2} \mathbb{E}\left[\int_a^b A_{i_1 j_1}(u)\, dW_u^{j_1} \int_a^b A_{i_2 j_1}(u)\, dW_u^{j_1} \int_a^b A_{i_1 j_2}(u)\, dW_u^{j_2} \int_a^b A_{i_2 j_2}(u)\, dW_u^{j_2}\right] = \mathbb{E}\left[\int_a^b A_{i_1 j_1}(u)\, dW_u^{j_1} \int_a^b A_{i_2 j_1}(u)\, dW_u^{j_1}\right] \mathbb{E}\left[\int_a^b A_{i_1 j_2}(u)\, dW_u^{j_2} \int_a^b A_{i_2 j_2}(u)\, dW_u^{j_2}\right],$$

Because of the properties of the stochastic integral, $\int_a^b A_{i_1 j}(u)\, dW_u^j = \mathcal{N}(0, \left(\int_a^b A_{i_1 j}^2(u)\, du\right)^{1/2})$ hence, because of the properties of the Gaussian integrals

$$\mathbb{E}\left[\left(\int_a^b A_{i_1 j}(u)\, dW_u^j\right)^4\right] = k\left(\int_a^b A_{i_1 j}^2(u)\, du\right)^2,$$

for some positive constant k. Moreover

$$\mathbb{E}\left[\left(\int_a^b A_{i_1 j_1}(u)\, dW_u^{j_1}\right)^2\right] = \int_a^b A_{i_1 j_1}^2(u)\, du,$$

and

$$\mathbb{E}\left[\int_a^b A_{i_1 j_1}(u)\, dW_u^{j_1} \int_a^b A_{i_2 j_1}(u)\, dW_u^{j_1}\right] = \int_a^b A_{i_1 j_1}(u) A_{i_2 j_1}(u)\, du.$$

This shows that the two terms $\mathbb{E}\left[\left\|\int_{t_0}^{s}(\Phi_L(t,s)-Id)\Phi_L(s,u)\mathbf{\Lambda}(u)\,d\mathbf{W}_u\right\|^4\right]$ and $\mathbb{E}\left[\left\|\int_{s}^{t}\Phi_L(t,u)\mathbf{\Lambda}(u)\,d\mathbf{W}_u\right\|^4\right]$ in (22) are of the order of $(t-s)^{1+a}$ where $a\geq 1$. Therefore we have

$$\mathbb{E}\left[\|F_1^{(n)}(X)_t - F_1^{(n)}(X)_s\|^4\right] \leq C|t-s|^{1+a}, \quad a\geq 1$$

for all s, t in $[t_0,T]$, where C is a constant independent of $t,\ s$. According to Kolmogorov criterion for tightness, the sequence of processes $\left\{F_1^{(n)}(X)\right\}_{n=0}^{\infty}$ is uniformly tight.

The proof for F_k, $k>1$ is similar. □

Let us note $\mu^n(t)$ (respectively $C^n(t,s)$) the mean (respectively the covariance matrix) function of $X_n = F_k(X_{n-1})$, $n\geq 1$. We have:

$$C^{n+1}(t,s) = \Phi_L(t,t_0)\Sigma^{Z_0}\Phi_L(s,t_0)^T + \int_{t_0}^{t\wedge s}\Phi_L(t,u)\mathbf{\Lambda}(u)\mathbf{\Lambda}(u)^T\Phi_L(s,u)^T\,du + \int_{t_0}^{t}\int_{t_0}^{s}\Phi_L(t,u)\mathrm{cov}\left(\widetilde{\mathbf{U}}_u^{X_n},\widetilde{\mathbf{U}}_v^{X_n}\right)\Phi_L(s,v)^T\,du\,dv \quad (23)$$

Note that the $kP\times kP$ covariance matrix $\mathrm{cov}\left(\widetilde{\mathbf{U}}_u^{X_n},\widetilde{\mathbf{U}}_v^{X_n}\right)$ has only one nonzero $P\times P$ block:

$$\mathrm{cov}\left(\widetilde{\mathbf{U}}_u^{X_n},\widetilde{\mathbf{U}}_v^{X_n}\right)_{kk} = \mathrm{cov}\left(\mathbf{U}_u^{X_n}\cdot\mathbf{1},\mathbf{U}_v^{X_n}\cdot\mathbf{1}\right),$$

We have

$$\mathrm{cov}\left(\mathbf{U}_u^{X_n}\cdot\mathbf{1},\mathbf{U}_v^{X_n}\cdot\mathbf{1}\right) = \mathrm{diag}\left(\sum_{\beta}\sigma_{\alpha\beta}^2\mathbb{E}\left[S_{\alpha\beta}(X_{n\beta}(u))S_{\alpha\beta}(X_{n\beta}(v))\right]\right),$$

and

$$\mathbb{E}\left[S_{\alpha\beta}(X_n(u))S_{\alpha\beta}(X_n(v))\right] = \int_{\mathbb{R}^2} S_{\alpha\beta}\left(\frac{\sqrt{C^n_{\beta\beta}(u,u)C^n_{\beta\beta}(v,v)-C^n_{\beta\beta}(u,v)^2}}{\sqrt{C^n_{\beta\beta}(u,u)}}x + \frac{C^n_{\beta\beta}(u,v)}{\sqrt{C^n_{\beta\beta}(u,u)}}y + \mu^n_\beta(v)\right) S_{\alpha\beta}\left(y\sqrt{C^n_{\beta\beta}(u,u)}+\mu^n_\beta(u)\right)\,Dx\,Dy, \quad (24)$$

where

$$Dx = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\,dx.$$

Similarly we have

$$\mu^n(t) = \Phi_L(t,t_0)\mu^{Z_0} + \int_{t_0}^t \Phi_L(t,u)\left(\mathbb{E}\left[\widetilde{\mathbf{U}}_u^{X_n}\right] + \tilde{\mathbf{I}}(u)\right) du =$$

$$\Phi_L(t,t_0)\mu^{Z_0} +$$

$$\int_{t_0}^t \Phi_L(t,u)\left(\left[0_P^T,\cdots,0_P^T,\left[\sum_\beta \bar{J}_{\alpha\beta}\int_{\mathbb{R}} S_{\alpha\beta}\left(x\sqrt{C^n_{\beta\beta}(u,u)}+\mu^n_\beta(u)\right)\right]_{\alpha=1,\cdots,P}\right]^T + \tilde{\mathbf{I}}(u)\right) Dxdu$$

We require the following four lemmas.

Lemma 3.4. For all $\alpha = 1,\cdots,P$ and $n \geq 1$ the quantity $C^n_{\alpha\alpha}(s,s)C^n_{\alpha\alpha}(t,t) - C^n_{\alpha\alpha}(t,s)^2$ is lowerbounded by the positive symmetric function:

$$\theta(s,t) = |t-s|\lambda^2_{\min}\lambda^{\Sigma^{Z_0}}_{\min}\lambda^{\Gamma}_{\min},$$

where $\lambda_{\min}$ is the smallest singular value of the positive symmetric definite matrix $\Phi_L(t,t_0)\Phi_L(t,t_0)^T$ for $t \in [t_0,T]$, $\lambda^{\Sigma^{Z_0}}_{\min}$ is the smallest eigenvalue of the positive symmetric definite covariance matrix Σ^{Z_0}, and $\lambda^{\Gamma}_{\min}$ is the smallest singular value of the matrix $\mathbf{\Lambda}(u)$ for $u \in [t_0,T]$.

Proof. We use equation (23) which we rewrite as follows, using the group property of the resolvent Φ_L :

$$C^{n+1}(t,s) = \Phi_L(t,t_0)\left(\Sigma^{Z_0} + \int_{t_0}^{t\wedge s}\Phi_L(t_0,u)\mathbf{\Lambda}(u)\mathbf{\Lambda}(u)^T\Phi_L(t_0,u)^T\,du + \right.$$

$$\left.\int_{t_0}^t\int_{t_0}^s \Phi_L(t_0,u)\mathrm{cov}\left(\widetilde{\mathbf{U}}_u^{X_n},\widetilde{\mathbf{U}}_v^{X_n}\right)\Phi_L(t_0,v)^T\,du\,dv\right)\Phi_L(t_0,s)^T.$$

We now assume $s<t$ and introduce the following notations, dropping the index n for simplicity:

$$\begin{array}{rcl} A(s) & = & \Sigma^{Z_0} + \int_{t_0}^s \Phi_L(t_0,u)\mathbf{\Lambda}(u)\mathbf{\Lambda}(u)^T\Phi_L(t_0,u)^T\,du \\ B(s,t) & = & \int_s^t \Phi_L(t_0,u)\mathbf{\Lambda}(u)\mathbf{\Lambda}(u)^T\Phi_L(t_0,u)^T\,du \\ a(t,s) & = & \int_{t_0}^t\int_{t_0}^s \Phi_L(t_0,u)\mathrm{cov}\left(\widetilde{\mathbf{U}}_u^{X_n},\widetilde{\mathbf{U}}_v^{X_n}\right)\Phi_L(t_0,v)^T\,du\,dv \end{array}$$

Let e_α, $\alpha = 1,\cdots,kP$, be the unit vector of the canonical basis whose coordinates are all equal to 0 except the αth one which is equal to 1. We note $E_\alpha(t)$ the vector $\Phi_L(t,t_0)^T e_\alpha$. We have

$$\begin{array}{rcl} C_{\alpha\alpha}(t,s) & = & E_\alpha(t)^T\left(A(s)+a(t,s)\right)E_\alpha(s) \\ C_{\alpha\alpha}(s,s) & = & E_\alpha(s)^T(A(s)+a(s,s))E_\alpha(s) \\ C_{\alpha\alpha}(t,t) & = & E_\alpha(t)^T(A(s)+B(s,t)+a(t,t))E_\alpha(t). \end{array}$$

Note that the last expression does not depend on s, since $A(s)+B(s,t)=A(t)$, which is consistent with the first equality. The reason why we introduce s in this expression is to simplify the following calculations.

The expression $C_{\alpha\alpha}(s,s)C_{\alpha\alpha}(t,t) - C_{\alpha\alpha}(t,s)^2$ is the sum of four sub-expressions:

$$e_1(s,t) = \left(E_\alpha(s)^T A(s) E_\alpha(s)\right)\left(E_\alpha(t)^T A(s) E_\alpha(t)\right) - \left(E_\alpha(t)^T A(s) E_\alpha(s)\right)^2,$$

which is greater than or equal to 0 because $A(s)$ is a covariance matrix,

$$e_2(s,t) = \left(E_\alpha(s)^T a(s,s) E_\alpha(s)\right)\left(E_\alpha(t)^T a(t,t) E_\alpha(t)\right) - \left(E_\alpha(t)^T a(t,s) E_\alpha(s)\right)^2,$$

which is also greater than or equal to 0 because $a(t,s)$ is a covariance matrix function,

$$\begin{aligned} e_3(s,t) = \left(E_\alpha(s)^T A(s) E_\alpha(s)\right)\left(E_\alpha(t)^T a(t,t) E_\alpha(t)\right) + \\ \left(E_\alpha(t)^T A(s) E_\alpha(t)\right)\left(E_\alpha(s)^T a(s,s) E_\alpha(s)\right) - \\ 2\left(E_\alpha(t)^T A(s) E_\alpha(s)\right)\left(E_\alpha(t)^T a(t,s) E_\alpha(s)\right) \end{aligned}$$

Because $a(t,s)$ is a covariance matrix function we have

$$E_\alpha(t)^T a(t,t) E_\alpha(t) + E_\alpha(s)^T a(s,s) E_\alpha(s) - 2E_\alpha(t)^T a(t,s) E_\alpha(s) \geq 0,$$

and , as seen above, $e_2(s,t) \geq 0$. Because $e_1(s,t) \geq 0$ we also have

$$\begin{aligned} -\sqrt{E_\alpha(s)^T A(s) E_\alpha(s)}\sqrt{E_\alpha(t)^T A(s) E_\alpha(t)} \leq E_\alpha(t)^T A(s) E_\alpha(s) \leq \\ \sqrt{E_\alpha(s)^T A(s) E_\alpha(s)}\sqrt{E_\alpha(t)^T A(s) E_\alpha(t)}, \end{aligned}$$

and, as it can be readily verified, this implies $e_3(s,t) \geq 0$.

Therefore we can lowerbound $C_{\alpha\alpha}(s,s)C_{\alpha\alpha}(t,t) - C_{\alpha\alpha}(t,s)^2$ by the fourth subexpression:

$$\begin{aligned} C_{\alpha\alpha}(s,s)C_{\alpha\alpha}(t,t) - C_{\alpha\alpha}(t,s)^2 \geq \left(E_\alpha(s)^T A(s) E_\alpha(s)\right)\left(E_\alpha(t)^T B(s,t) E_\alpha(t)\right) + \\ \left(E_\alpha(s)^T a(s,s) E_\alpha(s)\right)\left(E_\alpha(t)^T B(s,t) E_\alpha(t)\right) \geq \\ \left(E_\alpha(s)^T A(s) E_\alpha(s)\right)\left(E_\alpha(t)^T B(s,t) E_\alpha(t)\right), \end{aligned}$$

since $B(s,t)$ and $a(s,s)$ are covariance matrixes. We next have

$$E_\alpha(s)^T A(s) E_\alpha(s) = \frac{E_\alpha(s)^T A(s) E_\alpha(s)}{E_\alpha(s)^T E_\alpha(s)} \frac{e_\alpha^T \Phi_L(s,t_0)\Phi_L(s,t_0)^T e_\alpha}{e_\alpha^T e_\alpha},$$

by definition of $E_\alpha(s)$. Therefore

$$E_\alpha(s)^T A(s) E_\alpha(s) \geq \lambda_{\min}^{A(s)} \lambda_{\min}^{\Phi_L(s,t_0)\Phi_L(s,t_0)^T} \geq \lambda_{\min}^{\Sigma^{Z_0}} \lambda_{\min},$$

where $\lambda_{\min}^C$ is the smallest eigenvalue of the symmetric positive matrix C. Similarly we have

$$E_\alpha(t)^T B(s,t) E_\alpha(t) \geq \lambda_{\min}^{B(s,t)} \lambda_{\min}.$$

Let us write $\Gamma(u) = \mathbf{\Lambda}(u)\mathbf{\Lambda}(u)^T$. We have

$$\lambda_{\min}^{B(s,t)} = \min_{\|x\|\leq 1} \int_s^t \frac{x^T \Phi_L(t_0,u)\Gamma(u)\Phi_L(t_0,u)^T x}{x^T x}\, du = \min_{\|x\|\leq 1} \int_s^t \frac{x^T \Phi_L(t_0,u)\Gamma(u)\Phi_L(t_0,u)^T x}{x^T \Phi_L(t_0,u)\Phi_L(t_0,u)x} \frac{x^T \Phi_L(t_0,u)\Phi_L(t_0,u)x}{x^T x}\, du \geq \int_s^t \min_{\|x\|\leq 1} \left(\frac{x^T \Phi_L(t_0,u)\Gamma(u)\Phi_L(t_0,u)^T x}{x^T \Phi_L(t_0,u)\Phi_L(t_0,u)x} \frac{x^T \Phi_L(t_0,u)\Phi_L(t_0,u)x}{x^T x} \right) du \geq (t-s)\lambda_{\min}\lambda_{\min}^{\Gamma}.$$

Combining these results we have

$$C_{\alpha\alpha}(s,s)C_{\alpha\alpha}(t,t) - C_{\alpha\alpha}(t,s)^2 \geq |t-s|\lambda_{\min}^2 \lambda_{\min}^{\Sigma^{Z_0}} \lambda_{\min}^{\Gamma}$$

□

Lemma 3.5. For all $t \in [t_0, T]$ all $\alpha = 1, \cdots, P$, and $n \geq 1$, we have

$$C_{\alpha\alpha}^n(t,t) \geq k_0 > 0.$$

Proof. $C_{\alpha\alpha}^n(t,t)$ is larger than $(\Phi_L(t,t_0)\Sigma^{Z_0}\Phi_L(t,t_0)^T)_{\alpha\alpha}$ which is larger than the smallest eigenvalue of the matrix $\Phi_L(t,t_0)\Sigma^{Z_0}\Phi_L(t,t_0)^T$. This smallest eigenvalue is equal to

$$\min_x \frac{x^T \Phi_L(t,t_0)\Sigma^{Z_0}\Phi_L(t,t_0)^T x}{x^T x} = \min_x \frac{x^T \Phi_L(t,t_0)\Sigma^{Z_0}\Phi_L(t,t_0)^T x}{x^T \Phi_L(t,t_0)\Phi_L(t,t_0)^T x} \frac{x^T \Phi_L(t,t_0)\Phi_L(t,t_0)^T x}{x^T x} \geq \min_x \frac{x^T \Phi_L(t,t_0)\Sigma^{Z_0}\Phi_L(t,t_0)^T x}{x^T \Phi_L(t,t_0)\Phi_L(t,t_0)^T x} \min_x \frac{x^T \Phi_L(t,t_0)\Phi_L(t,t_0)^T x}{x^T x}.$$

In the last expression the first term is larger than the smallest eigenvalue $\lambda_{\min}^{\Sigma^{Z_0}}$ of the matrix Σ^{Z_0} which is positive definite since we have assumed the Gaussian random variable Z_0 nondegenerate. The second term is equal to the smallest singular value $\lambda_{\min}$ of the matrix $\Phi_L(t,t_0)$ which is also strictly positive for all $t \in [t_0, T]$ by hypothesis, see appendix B, equation (47).

□

We also use the following lemma.

Lemma 3.6. The $2n$-dimensional integral

$$I_n = \int_{[t_0, t\vee s]^2} \rho_1(u_1,v_1) \left(\int_{[t_0, u_1\vee v_1]^2} \cdots \left(\int_{[t_0, u_{n-2}\vee v_{n-2}]^2} \rho_{n-1}(u_{n-1}, v_{n-1}) \left(\int_{[t_0, u_{n-1}\vee v_{n-1}]^2} \rho_n(u_n,v_n) du_n dv_n \right) du_{n-1} dv_{n-1} \right) \cdots \right) du_1 dv_1,$$

where the functions $\rho_i(u_i, v_i)$, $i = 1, \cdots, n$ are either equal to 1 or to $1/\sqrt{\theta(u_i, v_i)}$, is upperbounded by $k^n/(n-1)!$ for some positive constant k.

Proof. First note that the integral is well-defined because of lemma 3.4. Second, note that there exists a constant K such that $K/\sqrt{\theta(u,v)} \geq 1$ for all $(u,v) \in [t_0, t\vee s]^2$, i.e. $K = \lambda_{\min}\sqrt{\lambda_{\min}^{\Sigma^{Z_0}}\lambda_{\min}^{\Gamma}(T-t_0)}$. Therefore the integral is upperbounded by K_0^n, where $K_0 = \max(1, K)$ times the integral obtained when $\rho_i(u_i, v_i) = 1/\sqrt{|u_i - v_i|}$ for all $i = 1, \cdots, n$. Let us then consider this situation. Without loss of generality we assume $t_0 = 0$. The cases $n = 1, 2, 3$ allow one to understand the process.

$$I_1 \leq K_0 \int_{[0,t\vee s]^2} \frac{dudv}{\sqrt{|u-v|}}. \tag{25}$$

Let us rotate the axes by $-\frac{\pi}{4}$ by performing the change of variables

$$u = \frac{U+V}{\sqrt{2}},$$
$$v = \frac{V-U}{\sqrt{2}}.$$

Using the symmetry of the integrand in s and t and the change of variable, the integral in the righthand side of (25) is equal to (see figure 2):

$$2\frac{1}{2^{1/4}}\int_0^{\frac{t\vee s}{\sqrt{2}}}\int_U^{\sqrt{2}(t\vee s)-U}\frac{dVdU}{\sqrt{U}} = 2^{3/4}\int_0^{a/2}\frac{a-2U}{\sqrt{U}}dU = 2^{3/4}\alpha_1 a^{3/2},$$

where $a = \sqrt{2}(t \vee s)$ and $\alpha_1 = \frac{2\sqrt{2}}{3}$.

Let us now look at I_2. It is upperbounded by the factor $K_0^2(2^{3/4})^2\alpha_1$ times the integral

$$\int_0^{a/2}\int_U^{a-U}\frac{(\sqrt{2}(u\vee v))^{3/2}}{\sqrt{U}}dUdV.$$

Since in the area of integration $u \vee v = v = \frac{V-U}{\sqrt{2}}$ we are led to the product of $2/5$ by the one-dimensional integral

$$\int_0^{a/2}\frac{(a-2U)^{5/2}}{\sqrt{U}}dUdV = \alpha_2 a^3,$$

where $\alpha_2 = \frac{5\sqrt{2}\pi}{32}$.

Similarly I_3 is upperbounded by the product of $K_0^3(2^{3/4})^3\alpha_1\alpha_2\frac{2}{5}\frac{2}{8}$ times the integral

$$\int_0^{a/2}\frac{(a-2U)^4}{\sqrt{U}}dUdV = \alpha_3 a^{9/2},$$

where $\alpha_3 = \frac{128\sqrt{2}}{315}$. One easily shows then that:

$$I_n \leq K_0^n F(2^{3/4})^n 2^n \left(\prod_{i=1}^n \alpha_i\right)\left(\frac{1}{\prod_{j=1}^n(2+3(j-1))}\right).$$

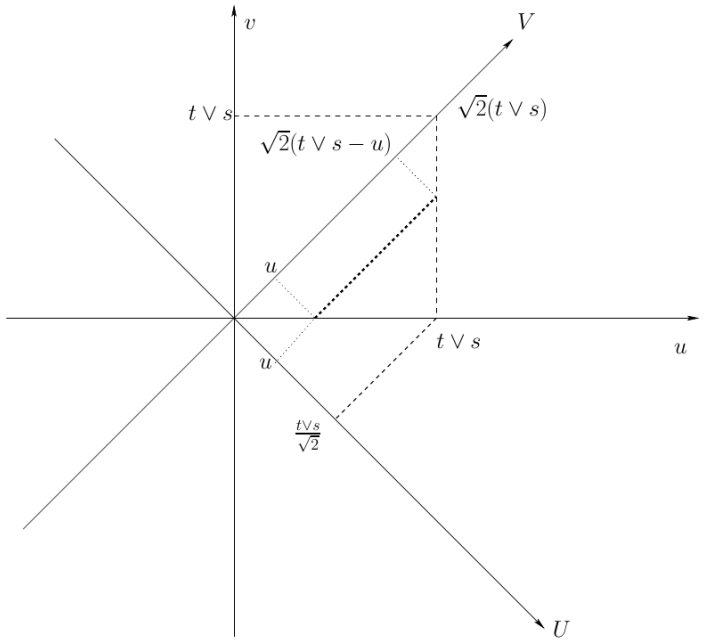


Figure 2: The change of coordinates.

It can be verified by using a system for symbolic computation that $0 < \alpha_i < 1$ for all $i \geq 1$. One also notices that

$$\prod_{j=1}^{n}(2+3(j-1)) \geq \frac{3^{n-1}}{2}(n-1)!,$$

therefore

$$I_n \leq K_0^n (2^{3/4})^n 2^{n-1} 3^{-(n-1)} \frac{1}{(n-1)!},$$

and this finishes the proof.

□

We now prove the following proposition.

Proposition 3.7. The sequences of covariance matrix functions $C^n(t,s)$ and of mean functions $\mu^n(t)$, s, t in $[t_0, T]$ are Cauchy sequences for the uniform norms.

Proof. We have

$$C^{n+1}(t,s) - C^n(t,s) = \int_{t_0}^{t}\int_{t_0}^{s} \Phi_L(t,u)\Big(\mathrm{cov}\left(\widetilde{\mathbf{U}}_u^{X_n}, \widetilde{\mathbf{U}}_v^{X_n}\right) - \mathrm{cov}\left(\widetilde{\mathbf{U}}_u^{X_{n-1}}, \widetilde{\mathbf{U}}_v^{X_{n-1}}\right)\Big)\Phi_L(s,v)^T\, du\, dv.$$

We take the infinite matrix norm of both sides of this equality and use the upperbounds $\|\Phi_L(t,u)\|_\infty \leq e^{\|\mathbf{L}\|_\infty (T-t_0)} = k_L$ and $\left\|\Phi_L(t,u)^T\right\|_\infty \leq e^{\left\|\mathbf{L}^T\right\|_\infty (T-t_0)} = k_{L^T}$ (see appendix B) to obtain

$$\begin{aligned}\left\|C^{n+1}(t,s) - C^n(t,s)\right\|_\infty &\leq k_L k_{L^T} \int_{t_0}^{t}\int_{t_0}^{s} \left\|\mathrm{cov}\left(\widetilde{\mathbf{U}}_u^{X_n}, \widetilde{\mathbf{U}}_v^{X_n}\right) - \mathrm{cov}\left(\widetilde{\mathbf{U}}_u^{X_{n-1}}, \widetilde{\mathbf{U}}_v^{X_{n-1}}\right)\right\|_\infty^v du\, dv \\ &= k_L k_{L^T} \int_{t_0}^{t}\int_{t_0}^{s} \left\|\mathrm{cov}\left(\mathbf{U}_u^{X_n}\cdot\mathbf{1}, \mathbf{U}_v^{X_n}\cdot\mathbf{1}\right) - \mathrm{cov}\left(\mathbf{U}_u^{X_{n-1}}\cdot\mathbf{1}, \mathbf{U}_v^{X_{n-1}}\cdot\mathbf{1}\right)\right\|_\infty^v du\, dv. \quad (26)\end{aligned}$$

According to equations (24) we are led to consider the difference $A_n - A_{n-1}$, where:

$$\begin{aligned}A_n &\stackrel{\mathrm{def}}{=} \\ &S_{\alpha\beta}\Big(\frac{\sqrt{C^n_{\beta\beta}(u,u)C^n_{\beta\beta}(v,v) - C^n_{\beta\beta}(u,v)^2}}{\sqrt{C^n_{\beta\beta}(u,u)}}x + \frac{C^n_{\beta\beta}(u,v)}{\sqrt{C^n_{\beta\beta}(u,u)}}y + \mu^n_\beta(v)\Big) S_{\alpha\beta}\left(y\sqrt{C^n_{\beta\beta}(u,u)} + \mu^n_\beta(u)\right) \\ &\stackrel{\mathrm{def}}{=} S_{\alpha\beta}\left[f^n_\beta(u,v)x + g^n_\beta(u,v)y + \mu^n_\beta(v)\right] S_{\alpha\beta}\left[h^n_\beta(u)y + \mu^n_\beta(u)\right].\end{aligned}$$

We write next:

$$A_n - A_{n-1} = S_{\alpha\beta}\left[f_\beta^n(u,v)x + g_\beta^n(u,v)y + \mu_\beta^n(v)\right] \left(S_{\alpha\beta}\left[h_\beta^n(u)y + \mu_\beta^n(u)\right] - S_{\alpha\beta}\left[h_\beta^{n-1}(u)y + \mu_\beta^{n-1}(u)\right]\right) + S_{\alpha\beta}\left[h_\beta^{n-1}(u)y + \mu_\beta^{n-1}(u)\right] \left(S_{\alpha\beta}\left[f_\beta^n(u,v)x + g_\beta^n(u,v)y + \mu_\beta^n(v)\right] - S_{\alpha\beta}\left[f_\beta^{n-1}(u,v)x + g_\beta^{n-1}(u,v)y + \mu_\beta^{n-1}(v)\right]\right).$$

The mean value theorem yields:

$$\mid A_n - A_{n-1} \mid \leq \|S_{\alpha\beta}\|_\infty \left\|S'_{\alpha\beta}\right\|_\infty \Big(\mid x \mid \mid f_\beta^n(u,v) - f_\beta^{n-1}(u,v) \mid + \mid y \mid \mid g_\beta^n(u,v) - g_\beta^{n-1}(u,v) \mid + \mid \mu_\beta^n(v) - \mu_\beta^{n-1}(v) \mid + \mid y \mid \mid h_\beta^n(u) - h_\beta^{n-1}(u) \mid + \mid \mu_\beta^n(u) - \mu_\beta^{n-1}(u) \mid \Big).$$

Using the fact that $\int_{-\infty}^{\infty} \mid x \mid \ Dx = \sqrt{\frac{2}{\pi}}$, we obtain:

$$\left\|C^{n+1}(t,s) - C^n(t,s)\right\|_\infty \leq k_L k_{L^T} k_C \Bigg(\sqrt{\frac{2}{\pi}} \int_{t_0}^t \int_{t_0}^s \left\|f^n(u,v) - f^{n-1}(u,v)\right\|_\infty du dv + \sqrt{\frac{2}{\pi}} \int_{t_0}^t \int_{t_0}^s \left\|g^n(u,v) - g^{n-1}(u,v)\right\|_\infty du dv + (t-t_0) \int_{t_0}^s \left\|\mu^n(v) - \mu^{n-1}(v)\right\|_\infty dv + (s-t_0) \int_{t_0}^t \left\|\mu^n(u) - \mu^{n-1}(u)\right\|_\infty du + \sqrt{\frac{2}{\pi}}(s-t_0) \int_{t_0}^t \left\|h^n(u) - h^{n-1}(u)\right\|_\infty du \Bigg),$$

where

$$k_C = \max_\alpha \sum_\beta \sigma_{\alpha\beta}^2 \|S_{\alpha\beta}\|_\infty \left\|S'_{\alpha\beta}\right\|_\infty . \tag{27}$$

A similar process applied to the mean values yields:

$$\left\|\mu^{n+1}(t) - \mu^n(t)\right\|_\infty \leq k_L \mu \Big(\int_{t_0}^t \left\|h^n(u) - h^{n-1}(u)\right\|_\infty du + \int_{t_0}^t \left\|\mu^n(u) - \mu^{n-1}(u)\right\|_\infty du \Big).$$

We now use the mean value theorem and lemmas 3.5 and 3.4 to find upperbounds for $\left\|f^n(u,v)-f^{n-1}(u,v)\right\|_\infty$, $\left\|g^n(u,v)-g^{n-1}(u,v)\right\|_\infty$ and $\left\|h^n(u)-h^{n-1}(u)\right\|_\infty$. We have

$$|h^n_\beta(u)-h^{n-1}_\beta(u)| = \left|\sqrt{C^n_{\beta\beta}(u,u)}-\sqrt{C^{n-1}_{\beta\beta}(u,u)}\right| \leq \frac{1}{2\sqrt{k_0}}\left|C^n_{\beta\beta}(u,u)-C^{n-1}_{\beta\beta}(u,u)\right|,$$

where k_0 is defined in lemma 3.5. Hence:

$$\left\|h^n(u)-h^{n-1}(u)\right\|_\infty \leq \frac{1}{2\sqrt{k_0}}\left\|C^n(u,u)-C^{n-1}(u,u)\right\|_\infty.$$

Along the same lines we can show easily that:

$$\left\|g^n(u,v)-g^{n-1}(u,v)\right\|_\infty \leq k\Big(\left\|C^n(u,v)-C^{n-1}(u,v)\right\|_\infty + \left\|C^n(u,u)-C^{n-1}(u,u)\right\|_\infty\Big),$$

and that:

$$\left\|f^n(u,v)-f^{n-1}(u,v)\right\|_\infty \leq \frac{k}{\sqrt{\theta(u,v)}}\Big(\left\|C^n(u,v)-C^{n-1}(u,v)\right\|_\infty + \left\|C^n(u,u)-C^{n-1}(u,u)\right\|_\infty + \left\|C^n(v,v)-C^{n-1}(v,v)\right\|_\infty\Big),$$

where $\theta(u,v)$ is defined in lemma 3.4. Grouping terms together and using the fact that all integrated functions are positive, we write:

$$\begin{aligned}\left\|C^{n+1}(t,s)-C^n(t,s)\right\|_\infty \leq \\ k\Bigg(&\int_{[t_0,t\vee s]^2}\frac{1}{\sqrt{\theta(u,v)}}\left\|C^n(u,v)-C^{n-1}(u,v)\right\|_\infty dudv+\\ &\int_{[t_0,t\vee s]^2}\frac{1}{\sqrt{\theta(u,v)}}\left\|C^n(u,u)-C^{n-1}(u,u)\right\|_\infty dudv+\\ &\int_{[t_0,t\vee s]^2}\left\|C^n(u,v)-C^{n-1}(u,v)\right\|_\infty dudv+\\ &\int_{[t_0,t\vee s]^2}\left\|C^n(u,u)-C^{n-1}(u,u)\right\|_\infty dudv+\\ &\int_{[t_0,t\vee s]^2}\left\|\mu^n(u)-\mu^{n-1}(u)\right\|_\infty dudv\Bigg). \quad (28)\end{aligned}$$

Note that, because of lemma 3.5, all integrals are well-defined. Regarding the mean functions, we write:

$$\left\|\mu^{n+1}(t)-\mu^n(t)\right\|_\infty \leq k\Bigg(\int_{[t_0,t\vee s]^2}\left\|C^n(u,u)-C^{n-1}(u,u)\right\|_\infty dudv+ \int_{[t_0,t\vee s]^2}\left\|\mu^n(u)-\mu^{n-1}(u)\right\|_\infty dudv\Bigg). \quad (29)$$

Proceeding recursively until we reach C^0 and μ^0 we obtain an upperbound for $\left\|C^{n+1}(t,s)-C^n(t,s)\right\|_\infty$ (respectively for $\left\|\mu^{n+1}(t)-\mu^n(t)\right\|_\infty$) which is the sum of less than 5^n terms each one being the product of k raised to a power less than or equal to n, times 2μ or 2Σ (upperbounds for the norms of the mean vector and the covariance matrix), times a $2n$-dimensional integral I_n given by

$$\int_{[t_0,t\vee s]^2}\rho_1(u_1,v_1)\Bigg(\int_{[t_0,u_1\vee v_1]^2}\cdots\Bigg(\int_{[t_0,u_{n-2}\vee v_{n-2}]^2}\rho_{n-1}(u_{n-1},v_{n-1}) \Bigg(\int_{[t_0,u_{n-1}\vee v_{n-1}]^2}\rho_n(u_n,v_n)du_ndv_n\Bigg)du_{n-1}dv_{n-1}\Bigg)\cdots\Bigg)du_1dv_1,$$

where the functions $\rho_i(u_i,v_i)$, $i=1,\cdots,n$ are either equal to 1 or to $1/\sqrt{\theta(u_i,v_i)}$. According to lemma 3.6, this integral is of the order of some positive constant raised to the power n divided by $(n-1)!$. Hence the sum is less than some positive constant k raised to the power n divided by $(n-1)!$. By taking the supremum with respect to t and s in $[t_0,T]$ we obtain the same result for $\left\|C^{n+1}-C^n\right\|_\infty$ (respectively for $\left\|\mu^{n+1}-\mu^n\right\|_\infty$). Since the series $\sum_{n\geq 1}\frac{k^n}{n!}$ is convergent, this implies that $\|C^{n+p}-C^n\|_\infty$ (respectively $\|\mu^{n+p}-\mu^n\|_\infty$) can be made arbitrarily small for large n and p and the sequence C^n (respectively μ^n) is a Cauchy sequence.

□

We can now prove the following theorem

Theorem 3.8. *For any nondegenerate kP-dimensional Gaussian random variable Z_0 and any initial process X such that $X(t_0)=Z_0$, the map F_k has a unique fixed point in $\mathcal{M}_1^+(C([t_0,T],\mathbb{R}^{kP}))$ towards which the sequence $\{F_k^{(n)}(X)\}_{n=0}^\infty$ of Gaussian processes converges in law.*

Proof. Since $C([t_0,T],\mathbb{R}^{kP})$ (respectively $C([t_0,T]^2,\mathbb{R}^{kP\times kP})$) is a Banach space for the uniform norm, the Cauchy sequence μ^n (respectively C^n) of proposition 3.7 converges to an element μ of $C([t_0,T],\mathbb{R}^{kP})$ (respectively an element C of $C([t_0,T]^2,\mathbb{R}^{kP\times kP})$). Therefore, according to theorem 3.1, the sequence $\{F_k^{(n)}(X)\}_{n=0}^\infty$ of Gaussian processes converges in law toward the Gaussian process Y with mean function μ and covariance function C. This process is clearly a fixed point of F_k.

Hence we know that there there exists at least one fixed point for the map F_k. Assume there exist two distinct fixed points Y_1 and Y_2 of F_k with mean functions μ_i and covariance functions

C_i, $i = 1, 2$, with the same initial condition. Since for all $n \geq 1$ we have $F_k^{(n)}(Y_i) = Y_i$, $i = 1, 2$, the proof of proposition 3.7 shows that $\|\mu_1^n - \mu_2^n\|_\infty$ (respectively $\left\|C_1^n - C_n^2\right\|_\infty$) is upper-bounded by the product of a positive number a_n (respectively b_n) with $\|\mu_1 - \mu_2\|_\infty$ (respectively with $\|C_1 - C_2\|_\infty$). Since $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$ and $\mu_i^n = \mu_i$, $i = 1, 2$ (respectively $C_i^n = C_i$, $i = 1, 2$), this shows that $\mu_1 = \mu_2$ and $C_1 = C_2$, hence the two Gaussian processes Y_1 and Y_2 are indistinguishable.

□

Conclusion

We have proved that for any non degenerate initial condition Z_0 there exists a unique solution of the mean field equations. The proof of theorem 3.8 is constructive, and hence provides a way for computing the solution of the mean field equations by iterating the map F_k, starting from any initial process X satisfying $X(t_0) = Z_0$, for instance a Gaussian process such as an Ornstein-Uhlenbeck process. We build upon these facts in section 5.

Note that the existence and uniqueness is true whatever the initial time t_0 and the final time T.

4 Existence and uniqueness of stationary solutions

So far, we have investigated the existence and uniqueness of solutions of the mean field equation for a given initial condition. We are now interested in investigating stationary solutions, which allow for some simplifications of the formalism.

A stationary solution is a solution whose probability distribution does not change under the flow of the equation. These solutions have been already investigated by several authors (see (Brunel and Hakim, 1999; Sompolinsky et al, 1988)). We propose a new framework to study and simulate these processes. Indeed we show in this section that under a certain contraction condition there exists a unique solution to the stationary problem. As in the previous section our proof is constructive and provides a way to simulate the solutions.

Remark. The long-time mean field description of a network is still a great endeavor in mathematics and statistical physics. In this section we formally take the mean field equation we obtained and let $t_0 \to -\infty$. This way we obtain an equation which is the limit of the mean field equation when $t_0 \to -\infty$. It means that we consider first the limit $N \to \infty$ and then $t_0 \to -\infty$. These two limits do not necessarily commute and there are known examples, for instance in spin glasses, where they do not.

It is clear that in order to get stationary solutions, we have to assume that the leak matrix $\mathbf{L}(t)$ does not depend upon t. Therefore, the resolvent $\Phi_L(t, s)$ is equal to $e^{\mathbf{L}(t-s)}$. To ensure stability of the solutions and the existence of a stationary process we also assume that the real parts of its eigenvalues are negative:

$$\mathrm{Re}(\lambda) < -\lambda_L \quad \lambda_L > 0 \tag{30}$$

for all eigenvalues λ of $\mathbf{L}$. This implies that we only consider first-order system since otherwise the matrix $\mathbf{L}$ has eigenvalues equal to 0.

For the same reason, we assume that the noise matrix $\mathbf{\Lambda}(t)$ and the input currents $\mathbf{I}(t)$ are constant in time. We further assume that the matrix $\mathbf{\Lambda}$ has full rank.

Proposition 4.1. Under the previous assumptions we have

$$\begin{cases} \lim\limits_{t_0 \to -\infty} e^{\mathbf{L}(t-t_0)} = 0, \\ \int_{-\infty}^{t} \left\| e^{\mathbf{L}(t-s)} \right\| \, ds = \int_0^{\infty} \left\| e^{\mathbf{L}u} \right\|_{\infty} \, du \stackrel{\text{def}}{=} M_L < \infty, \\ \int_{-\infty}^{t} \left\| e^{\mathbf{L}^T(t-s)} \right\|_{\infty} \, ds = \int_0^{\infty} \left\| e^{\mathbf{L}^T u} \right\|_{\infty} \, du \stackrel{\text{def}}{=} M_{L^T} < \infty, \end{cases}$$

and the process $Y_t^{t_0} = \int_{t_0}^{t} e^{\mathbf{L}(t-s)} \mathbf{\Lambda} \cdot d\mathbf{W}_s$ is well-defined, Gaussian and stationary when $t_0 \to -\infty$.

Proof. The first property follows from the fact that $\mathrm{Re}(\lambda) < -\lambda_L$ for all eigenvalues λ of $\mathbf{L}$. This assumption also implies that there exists a norm on $\mathbb{R}^P$ such that

$$\left\| e^{\mathbf{L}t} \right\| \leq e^{-\lambda_L t} \; \forall t \geq 0,$$

and hence

$$\left\| e^{\mathbf{L}t} \right\|_{\infty} \leq k e^{-\lambda_L t} \; \forall t \geq 0, \tag{31}$$

for some positive constant k. This implies the remaining two properties.

The stochastic integral $\int_{t_0}^{t} e^{\mathbf{L}(t-s)} \mathbf{\Lambda} \cdot d\mathbf{W}_s$ is well-defined $\forall t \leq T$ and is Gaussian with zero-mean. We note $Y_t^{t_0}$ the corresponding process. Its covariance matrix reads:

$$\Sigma^{Y_t^{t_0} Y_{t'}^{t_0}} = \int_{t_0}^{t \wedge t'} e^{\mathbf{L}(t-s)} \mathbf{\Lambda} \mathbf{\Lambda}^T e^{\mathbf{L}^T(t'-s)} \, ds.$$

Let us assume for instance that $t' < t$ and perform the change of variable $u = t - s$ to obtain

$$\Sigma^{Y_t^{t_0} Y_{t'}^{t_0}} = \left(\int_{t-t'}^{t-t_0} e^{\mathbf{L}u} \mathbf{\Lambda} \mathbf{\Lambda}^T e^{\mathbf{L}^T u} \, du \right) e^{\mathbf{L}^T(t'-t)}.$$

Under the previous assumptions this matrix integral is defined when $t_0 \to -\infty$ (dominated convergence theorem) and we have

$$\Sigma^{Y_t^{-\infty} Y_{t'}^{-\infty}} = \left(\int_{t-t'}^{+\infty} e^{\mathbf{L}u} \mathbf{\Lambda} \mathbf{\Lambda}^T e^{\mathbf{L}^T u} \, du \right) e^{\mathbf{L}^T(t'-t)},$$

which is a function of $t' - t$. □

This guarantees that there exists a stationary distribution of the equation:

$$d\mathbf{X}_0(t) = \mathbf{L} \cdot \mathbf{X}_0(t) \, dt + \mathbf{\Lambda} \cdot d\mathbf{W}_t, \tag{32}$$

such that $\mathbb{E}\left[\mathbf{X}_0(t)\right] = 0$. We have

$$\mathbf{X}_0(t) = \int_{-\infty}^{t} e^{\mathbf{L}(t-s)} \mathbf{\Lambda} \cdot d\mathbf{W}_s.$$

Its covariance matrix Σ^0 is equal to $\Sigma^{Y_t^{-\infty}Y_t^{-\infty}}$ and is independent of t.

We call *long term mean field equation* (LTMFE) the implicit equation:

$$\mathbf{V}(t) = \int_{-\infty}^{t} e^{\mathbf{L}(t-s)}\Big(\mathbf{U}_s^{\mathbf{V}} \cdot \mathbf{1} + \mathbf{I}\Big)\, ds + \mathbf{X}_0(t) \tag{33}$$

where $\mathbf{X}_0$ is the stationary process defined by equation (32) and where $\mathbf{U}^{\mathbf{V}}(t)$ is the effective interaction process introduced previously.

We next define the long term function $F_{\text{stat}} : \mathcal{M}_1^+(C((-\infty,T],\mathbb{R}^P) \to \mathcal{M}_1^+(C((-\infty,T],\mathbb{R}^P)$:

$$F_{\text{stat}}(\mathbf{X})_t = \int_{-\infty}^{t} e^{\mathbf{L}(t-s)}\Big(\mathbf{U}_s^{X} \cdot \mathbf{1} + \mathbf{I}\Big)\, ds + \mathbf{X}_0(t).$$

Proposition 4.2. The function F_{stat} is well defined on $\mathcal{M}_1^+(C((-\infty,T],\mathbb{R}^P)$.

Proof. We have already seen that the process $\mathbf{X}_0$ is well defined. The term $\int_{-\infty}^{t} e^{\mathbf{L}(t-s)}\mathbf{I}\, ds = \left(\int_{-\infty}^{t} e^{\mathbf{L}(t-s)}\, ds\right)\mathbf{I}$ is also well defined because of the assumptions on $\mathbf{L}$.

Let X be a given process in $\mathcal{M}_1^+(C((-\infty,T],\mathbb{R}^P)$. To prove the proposition we just have to ensure that the Gaussian process $\int_{-\infty}^{t} e^{\mathbf{L}(t-s)}\mathbf{U}_s^X \cdot \mathbf{1}\, ds$ is well defined. This results from the contraction assumption on $\mathbf{L}$ and the fact that the functions $S_{\alpha\beta}$ are bounded. We decompose this process into a "long memory" term $\int_{-\infty}^{0} e^{\mathbf{L}(t-s)}\mathbf{U}_s^X \cdot \mathbf{1}\, ds$ and the interaction term from time $t=0$, namely $\int_0^t e^{\mathbf{L}(t-s)}\mathbf{U}_s^X \cdot \mathbf{1}\, ds$. This latter term is clearly well defined. We show that the memory term is also well defined as a Gaussian random variable.

We write this term $e^{\mathbf{L}t}\int_{-\infty}^{0} e^{-\mathbf{L}s}\mathbf{U}_s^X \cdot \mathbf{1}\, ds$ and consider the second factor. This random variable is Gaussian, its mean reads $\int_0^{\infty} e^{\mathbf{L}s}\mu^{\mathbf{U}_{-s}^X} \cdot \mathbf{1}\, ds$ where

$$\mu^{\mathbf{U}_{-s}^X} = \left(\sum_{\beta=1}^{P} \bar{J}_{\alpha\beta}\mathbb{E}\left[S_{\alpha\beta}(X_\beta(-s))\right] + I_\alpha\right)_{\alpha=1\ldots P}$$

The integral defining the mean is well-defined because of (31) and the fact that the functions $S_{\alpha\beta}$ are bounded. A similar reasoning shows that the corresponding covariance matrix is well-defined. Hence the Gaussian process $\int_{-\infty}^{0} e^{-\mathbf{L}s}\mathbf{U}_s^X \cdot \mathbf{1} ds$ is well defined, and hence for any process $X \in \mathcal{M}_1^+(C((-\infty,T],\mathbb{R}^P)$, the process $F_{\text{stat}}(X)$ is well defined. □

We can now prove the following proposition.

Proposition 4.3. The mean vectors and the covariance matrices of the processes in the image of F_{stat} are bounded.

Proof. Indeed, since $\mathbb{E}[X_0(t)] = 0$, we have:

$$\|\mathbb{E}[F_{\text{stat}}(\mathbf{X})_t]\|_\infty = \left\| \int_{-\infty}^{t} e^{\mathbf{L}(t-s)} \mu^{\mathbf{U}_s^X} ds \right\|_\infty \leq M_L(\mu + \|I\|_\infty) \stackrel{\text{def}}{=} \mu_{LT}.$$

In a similar fashion the covariance matrices of the processes in the image of F_{stat} are bounded. Indeed we have:

$$\mathbb{E}\left[F_{\text{stat}}(\mathbf{X})_t F_{\text{stat}}(\mathbf{X})_t^T\right] = \Sigma^0 + \int_{-\infty}^{t}\int_{-\infty}^{t} e^{\mathbf{L}(t-s_1)} \text{diag}\left(\sum_\beta \sigma_{\alpha\beta}^2 \mathbb{E}[S_{\alpha\beta}(X_\beta(s_1))S_{\alpha\beta}(X_\beta(s_2))]\right) e^{\mathbf{L}^T(t-s_2)}\, ds_1\, ds_2,$$

resulting in

$$\left\|\mathbb{E}\left[F_{\text{stat}}(\mathbf{X})_t F_{\text{stat}}(\mathbf{X})_t^T\right]\right\|_\infty \leq \left\|\Sigma^0\right\|_\infty + k^2 \left(\frac{\sigma_{\max}}{\lambda_L}\right)^2 \stackrel{\text{def}}{=} \Sigma_{LT}.$$

□

Lemma 4.4. The set of stationary processes is invariant by F_{stat}.

Proof. Since the processes in the image of F_{stat} are Gaussian processes, one just needs to check that the mean of the process is constant in time and that its covariance matrix $C(s,t)$ only depends on $t-s$.

Let Z be a stationary process and $Y = F_{\text{stat}}(Z)$. We denote by μ_α^Z the mean of the process $Z_\alpha(t)$ and by $C_\alpha^Z(t-s)$ its covariance function. The mean of the process $U_{\alpha\beta}^Z$ reads:

$$m_{\alpha,\beta}^Z(t) = \mathbb{E}[S_{\alpha\beta}(Z_\beta(t))] = \frac{1}{\sqrt{2\pi C_\beta^Z(0)}} \int_{\mathbb{R}} S_{\alpha\beta}(x) e^{\frac{(x-\mu_\beta^Z)^2}{2C_\beta^Z(0)}} dx$$

and hence does not depends on time. We note μ^Z the mean vector of the stationary process $\mathbf{U}^Z \cdot \mathbf{1}$.

Similarly, its covariance function reads:

$$\Delta_{\alpha\beta}^Z(t,s) = \mathbb{E}[S_{\alpha\beta}(Z_\beta(t))S_{\alpha\beta}(Z_\beta(s))] = \int_{\mathbb{R}^2} S_{\alpha\beta}(x)S_{\alpha\beta}(y) \exp\left(-\frac{1}{2}\begin{pmatrix} x-\mu_\beta^Z \\ y-\mu_\beta^Z \end{pmatrix}^T \begin{pmatrix} C_\beta^Z(0) & C_\beta^Z(t-s) \\ C_\beta^Z(t-s) & C_\beta^Z(0) \end{pmatrix}^{-1} \begin{pmatrix} x-\mu_\beta^Z \\ y-\mu_\beta^Z \end{pmatrix}\right) dx\, dy$$

which is clearly a function, noted $\Delta_{\alpha\beta}^Z(t-s)$, of $t-s$. Hence $\mathbf{U}^Z \cdot \mathbf{1}$ is stationary and we denote by $C^{U^Z}(t-s)$ its covariance function.

It follows that the mean of Y_t reads:

$$\begin{aligned}\mu^Y(t) &= \mathbb{E}\left[F_{\text{stat}}(Z)_t\right] \\ &= \mathbb{E}\left[X_0(t)\right] + \mathbb{E}\left[\int_{-\infty}^t e^{\mathbf{L}(t-s)}\left(\mathbf{I} + \mathbf{U}_s^Z \cdot \mathbf{1}\right)\, ds\right] \\ &= \int_{-\infty}^t e^{\mathbf{L}(t-s)}\left(\mathbf{I} + \mathbb{E}\left[\mathbf{U}_s^Z \cdot \mathbf{1}\right]\right)\, ds \\ &= \left(\int_{-\infty}^0 e^{\mathbf{L}u}\, du\right)\left(\mathbf{I} + \mu^Z\right)\end{aligned}$$

Since we proved that $\mathbb{E}\left[\mathbf{U}_s^Z \cdot \mathbf{1}\right] = \mu^Z$ was not a function of s.

Similarly, we compute the covariance function and check that it can be written as a function of $(t-s)$. Indeed, it reads:

$$\begin{aligned}C^Y(t,s) &= \int_{-\infty}^t \int_{-\infty}^s e^{\mathbf{L}(t-u)} Cov(\mathbf{U}_u^Z \cdot \mathbf{1}, \mathbf{U}_v^Z \cdot \mathbf{1}) e^{\mathbf{L}^T(s-v)}\, du\, dv + Cov(X_0(t), X_0(s)) \\ &= \int_{-\infty}^0 \int_{-\infty}^0 e^{\mathbf{L}u} C^{U^Z}(t-s+(u-v)) e^{\mathbf{L}^T v}\, du\, dv + C^{X_0}(t-s)\end{aligned}$$

since the process X_0 is stationary. $C^Y(t,s)$ is clearly a function of $t-s$. Hence Y is a stationary process, and the proposition is proved.

□

Theorem 4.5. *The sequence of processes $\{F_{stat}^{(n)}(X)\}_{n=0}^{\infty}$ is uniformly tight.*

Proof. The proof is essentially the same as the proof of theorem 3.3, since we can write

$$F_{\text{stat}}(X)_t = e^{\mathbf{L}t} F_{\text{stat}}(X)_0 + \int_0^t e^{\mathbf{L}(t-s)}(\mathbf{U}_s^X \cdot \mathbf{1} + \mathbf{I})\, ds + \int_0^t e^{\mathbf{L}(t-u)} \mathbf{\Lambda} d\mathbf{W}_s$$

$F_{\text{stat}}(X)_t$ appears as the sum of the random variable $F_{\text{stat}}(X)_0$ and the Gaussian process defined by $\int_0^t e^{\mathbf{L}(t-s)}(\mathbf{U}_s^X \cdot \mathbf{1} + \mathbf{I})\, ds + \int_0^t e^{\mathbf{L}(t-u)} \mathbf{\Lambda} d\mathbf{W}_s$ which is equal to $F_k(X)_t$ defined in section 3 for $t_0 = 0$. Therefore $F_{\text{stat}}^{(n)}(X)_t = F_k^{(n)}(X)_t$ for $t > 0$. We have proved the uniform tightness of the sequence of processes $\{F_k^{(n)}(X)\}_{n=0}^{\infty}$ in theorem 3.3. Hence, according to Kolmogorov's criterion for tightness, we just have to prove that the sequence of Gaussian random variables:

$$F_{\text{stat}}^{(n)}(X)_0 = \left\{\int_{-\infty}^0 \Phi_L(-u)(\mathbf{U}_u^{F_{\text{stat}}^{(n)}(X)} \cdot \mathbf{1} + \mathbf{I}) du + \mathbf{X}_0(0)\right\}_{n \geq 0}$$

is uniformly tight. Since it is a sequence of Gaussian random variables, it is sufficient to prove that their means and covariance matrices are upperbounded to obtain that for any $\varepsilon > 0$ there exists a compact K_ε such that for any $n \in \mathbb{N}$, we have $\mathbb{P}(F_{\text{stat}}^{(n)}(X)_0 \in K_\varepsilon) \geq 1 - \varepsilon$. This is a consequence

of proposition 4.3 for the first random variable and of the definition of $\mathbf{X}_0$ for the second. By Kolmogorov's criterion the sequence of processes $\{F^{(n)}_{\text{stat}}(X)\}_{n=0}^{\infty}$ is uniformly tight

□

In order to apply theorem 3.1 we need to prove that the sequences of covariance and mean functions are convergent. Unlike the case of t_0 finite, this is not always true. Indeed, to ensure existence and uniqueness of solutions in the stationary case, the parameters of the system have to satisfy a contraction condition, and proposition 3.7 extends as follows.

Proposition 4.6. If λ_L defined in (30) satisfies the conditions (34) defined in the proof, depending upon k_C (defined in (27)), k_0, μ_{LT} and Σ_{LT} (defined in proposition 4.3)then the sequences of covariance matrix functions $C^n(t,s)$ and of mean functions $\mu^n(t)$, $s,\ t$ in $[t_0,T]$ are Cauchy sequences for the uniform norms.

Proof. The proof follows that of proposition 3.7 with a few modifications that we indicate. In establishing the equation corresponding to (26) we use the fact that $\|\Phi_L(t,u)\|_\infty \leq ke^{-\lambda_L(t-u)}$ for some positive constant k and all u, t, $u \leq t$. We therefore have:

$$\left\|C^{n+1}(t,s)-C^n(t,s)\right\|_\infty \leq$$
$$k^2e^{-\lambda_L(t+s)}\int_{-\infty}^{t}\int_{-\infty}^{s}e^{\lambda_L(u+v)}\left\|Cov\left(\mathbf{U}_u^{X_n},\mathbf{U}_v^{X_n}\right)-Cov\left(\mathbf{U}_u^{X_{n-1}},\mathbf{U}_v^{X_{n-1}}\right)\right\|_\infty^v du\,dv$$

The rest of the proof proceeds the same way as in proposition 3.7. Equations (28) and (29) become:

$$\left\|C^{n+1}(t,s)-C^n(t,s)\right\|_\infty \leq$$
$$Ke^{-\lambda_L(t+s)}\left(\int_{[-\infty,t\vee s]^2}\frac{e^{\lambda_L(u+v)}}{\sqrt{f(u,v)}}\left\|C^n(u,v)-C^{n-1}(u,v)\right\|_\infty dudv+\right.$$
$$\int_{[-\infty,t\vee s]^2}\frac{e^{\lambda_L(u+v)}}{\sqrt{f(u,v)}}\left\|C^n(u,u)-C^{n-1}(u,u)\right\|_\infty dudv+$$
$$\int_{[-\infty,t\vee s]^2}e^{\lambda_L(u+v)}\left\|C^n(u,v)-C^{n-1}(u,v)\right\|_\infty dudv+$$
$$\int_{[-\infty,t\vee s]^2}e^{\lambda_L(u+v)}\left\|C^n(u,u)-C^{n-1}(u,u)\right\|_\infty dudv+$$
$$\left.\int_{[-\infty,t\vee s]^2}e^{\lambda_L(u+v)}\left\|\mu^n(u)-\mu^{n-1}(u)\right\|_\infty dudv\right),$$

and

$$\left\|\mu^{n+1}(t)-\mu^n(t)\right\|_\infty \leq Ke^{-\lambda_L(t+s)}\left(\int_{[-\infty,t\vee s]^2} e^{\lambda_L(u+v)}\left\|C^n(u,u)-C^{n-1}(u,u)\right\|_\infty dudv+\right.$$
$$\left.\int_{[-\infty,t\vee s]^2} e^{\lambda_L(u+v)}\left\|\mu^n(u)-\mu^{n-1}(u)\right\|_\infty dudv\right),$$

for some positive constant K, function of k, k_C (defined in (27)), and k_0.

Proceeding recursively until we reach C^0 and μ^0 we obtain an upperbound for $\left\|C^{n+1}(t,s)-C^n(t,s)\right\|_\infty$ (respectively for $\left\|\mu^{n+1}(t)-\mu^n(t)\right\|_\infty$) which is the sum of less than 5^n terms each one being the product of K^n, times $2\mu_{LT}$ or $2\Sigma_{LT}$, times a $2n$-dimensional integral I_n given by:

$$\int_{[-\infty,t\vee s]^2}\rho_1(u_1,v_1)\left(\int_{[-\infty,u_1\vee v_1]^2}\cdots\left(\int_{[-\infty,u_{n-2}\vee v_{n-2}]^2}\rho_{n-1}(u_{n-1},v_{n-1})\right.\right.$$
$$\left.\left.\left(\int_{[-\infty,u_{n-1}\vee v_{n-1}]^2} e^{\lambda_L(u_n+v_n)}\rho_n(u_n,v_n)du_ndv_n\right)du_{n-1}dv_{n-1}\right)\cdots\right)du_1dv_1,$$

where the functions $\rho_i(u_i,v_i)$, $i=1,\cdots,n$ are either equal to 1 or to $1/\sqrt{\theta(u_i,v_i)}$.

It can be shown by straightforward calculation that each sub-integral contributes at most either

$$\frac{K_0}{\lambda_L^2} \quad \text{if} \quad \rho_i=1 \quad \text{or} \quad \sqrt{\frac{\pi}{2}}\frac{K_0}{\lambda_L^{3/2}},$$

in the other case. Hence we obtain factors of the type

$$K_0^n\left(\frac{1}{\lambda_L^2}\right)^p\left(\sqrt{\frac{\pi}{2}}\frac{1}{\lambda_L^{3/2}}\right)^{n-p}=\left(\sqrt{\frac{\pi}{2}}\right)^{n-p}\left(\frac{1}{\lambda_L}\right)^{(3n+p)/2}K_0^n,$$

where $0\leq p\leq n$. If $\lambda_L<1$, $(\lambda_L)^{(3n+p)/2}\geq\lambda_L^{2n}$ and else $(\lambda_L)^{(3n+p)/2}\geq\lambda_L^{3n/2}$. Since $\left(\sqrt{\frac{\pi}{2}}\right)^{n-p}\leq\left(\sqrt{\frac{\pi}{2}}\right)^n$ we obtain the two conditions

$$1>\lambda_L^2\geq 5\sqrt{\frac{\pi}{2}}KK_0 \quad \text{or} \quad \left\{\lambda_L^{3/2}>5\sqrt{\frac{\pi}{2}}KK_0 \quad \text{and} \quad \lambda_L\geq 1\right\} \tag{34}$$

□

Putting all these results together we obtain the following theorem of existence and uniqueness of solutions for the long term mean field equations:

Theorem 4.7. *Under the contraction conditions* (34)*, the function F_{stat} has a unique solution in $\mathcal{M}_1^+(C((-\infty,T],\mathbb{R}^P)$ which is stationary, and for any process X, the sequence $\{F_{stat}^{(n)}(X)\}_{n=0}^\infty$ of Gaussian processes converges in law toward the unique fixed point of the function F_{stat}.*

Proof. The proof is essentially similar to the one of theorem 3.8. Indeed, the mean and the covariance matrixes converge since they are Cauchy sequences in the complete space of continuous functions equipped with the uniform norm. Using theorem 3.1, we obtain that the sequence converges to a process Y which is necessarily a fixed point of F_{stat}. Hence we have existence of a fixed point for F_{stat}. The uniqueness comes from the results obtained in the proof of proposition 4.6. The limiting process is necessarily stationary. Indeed, let X be a stationary process. Then for any $n \in \mathbb{N}$, the process $F_{\text{stat}}^{(n)}(X)$ will be stationary by the virtue of lemma 4.4, and hence so will be the limiting process which is the only fixed point of F_{stat}.

□

Hence in the stationary case, the existence and uniqueness of a solution is not always ensured. For instance if the leaks are too small (i.e. when the time constants of the decay of the membrane potentials are too long) then the sequence can diverge or have multiple fixed points.

5 Numerical experiments

5.1 Simulation algorithm

Beyond the mathematical results, the framework that we introduced in the previous sections gives us a strategy to compute numerically the solutions of the dynamic mean-field equations. Indeed, we proved in section 3 that under very moderate assumptions on the covariance matrix of the noise, the iterations of the map F_k starting from any initial condition converge to the solution of the mean field equations.

This convergence result gives us a direct way to compute numerically the solution of the mean field equations. Since we are dealing with Gaussian processes, determining the law of the iterates of the map F_k amounts to computing its mean and covariance functions. In this section we describe our numerical algorithm in the case of the Simple Model of section 2.2.2.

5.1.1 Computing F_k.

Let X be a P-dimensional Gaussian process of mean $\mu^X = (\mu_\alpha^X(t))_{\alpha=1\dots P}$ and covariance $C^X = (C_{\alpha\beta}^X(s,t))_{\alpha,\beta\in\{1\dots P\}}$. We fix a time interval $[t_0 = 0, T]$ and denote by Y the image of the process X under F_1. In the case of the simple model, the covariance of Y is diagonal. Hence in this case the expressions we obtain in section 3 simply read:

$$\begin{aligned}\mu_\alpha^Y(t) &= \mu_\alpha^X(0)e^{-t/\tau_\alpha} + \int_0^t e^{-(t-s)/\tau_\alpha}\left(\sum_{\beta=1}^P \bar{J}_{\alpha\beta}\mathbb{E}\left[S_{\alpha,\beta}(X_\beta(s))\right] + I_\alpha(s)\right)ds \\ &= \mu_\alpha^X(0)e^{-t/\tau_\alpha} + \int_0^t e^{-(t-s)/\tau_\alpha} I_\alpha(s)ds \\ &\quad + \sum_{\beta=1}^P \bar{J}_{\alpha\beta}\int_0^t e^{-(t-s)/\tau_\alpha}\int_{-\infty}^{+\infty} S_{\alpha\beta}\left(x\sqrt{v_\beta^X(s)} + \mu_\beta^X(s)\right)Dxds.\end{aligned}$$

where we denoted $v_\alpha^X(s)$ the standard deviation of X_α at time s, instead of $C_{\alpha\alpha}^X(s,s)$. Thus, knowing $v_\alpha^X(s), s \in [0,t]$ we can compute $\mu_\alpha^Y(t)$ using a standard discretization scheme of the integral, with

a small time step compared with τ_α and the characteristic time of variation of the input current I_α. Alternatively, we can use the fact that μ_α^Y satisfies the differential equation:

$$\frac{d\mu_\alpha^Y}{dt} = -\frac{\mu_\alpha^Y}{\tau_\alpha} + \sum_{\beta=1}^{P} \bar{J}_{\alpha\beta} \int_{-\infty}^{+\infty} S_{\alpha\beta}\left(x\sqrt{v_\beta^X(t)} + \mu_\beta^X(t)\right) Dx + I_\alpha(t),$$

and compute the solution using a Runge-Kutta algorithm (which is faster and more accurate). Note that, when all the standard deviations of the process X are null for all time $t \in [0,T]$, we obtain a standard dynamical system. Nevertheless, in the general case, $v_\beta^X(t) > 0$ for some βs, and the dynamical evolution of μ_α^Y depends on the Gaussian fluctuations of the field X. These fluctuations must be computed via the complete equation of the covariance diagonal coefficient $C_{\alpha\alpha}^Y(t,s)$, which reads:

$$C_{\alpha\alpha}^Y(t,s) = e^{-(t+s)/\tau_\alpha}\Big[v_\alpha^X(0) + \frac{\tau_\alpha s_\alpha^2}{2}\left(e^{\frac{2s}{\tau_\alpha}} - 1\right) + \sum_{\beta=1}^{P} \sigma_{\alpha\beta}^2 \int_0^t \int_0^s e^{(u+v)/\tau_\alpha} \Delta_{\alpha\beta}^X(u,v) du dv\Big],$$

where:

$$\Delta_{\alpha\beta}^X(u,v) = \int_{\mathbb{R}^2} S_{\alpha\beta}\left(x\frac{\sqrt{v_\beta^X(u)v_\beta^X(v) - C_{\beta\beta}^X(u,v)^2}}{\sqrt{v_\beta^X(v)}} + y\frac{C_{\beta\beta}^X(u,v)}{\sqrt{v_\beta^X(v)}} + \mu_\beta^X(u)\right) \times S_{\alpha\beta}\left(y\sqrt{v_\beta^X(v)} + \mu_\beta^X(v)\right) Dx\, Dy.$$

Unless if we assume the stationarity of the process (see e.g. section 5.2), this equation cannot be written as an ordinary differential equation. We clearly observe here the non-Markovian nature of the problem: $C_{\alpha\alpha}^X(t,s)$ depends on the whole past of the process until time $t \vee s$.

This covariance can be split into the sum of two terms: the external noise contribution $C_{\alpha\alpha}^{OU}(t,s) = e^{-(t+s)/\tau_\alpha}\left[v_\alpha^X(0) + \frac{\tau_\alpha s_\alpha^2}{2}\left(e^{\frac{2s}{\tau_\alpha}} - 1\right)\right]$ and the interaction between the neurons. The external noise contribution is a simple function and can be computed directly. To compute the interactions contribution to the standard deviation we have to compute the symmetric two-variables function:

$$H_{\alpha\beta}^X(t,s) = e^{-(t+s)/\tau_\alpha} \int_0^t \int_0^s e^{(u+v)/\tau_\alpha} \Delta_{\alpha\beta}^X(u,v) du dv,$$

from which one obtains the standard deviation using the formula

$$C_{\alpha\alpha}^Y(t,s) = C_{\alpha\alpha}^{OU}(t,s) + \sum_{\beta=1}^{P} \sigma_{\alpha\beta}^2 H_{\alpha\beta}^X(t,s).$$

To compute the function $H^X_{\alpha\beta}(t,s)$, we start from $t=0$ and $s=0$, where $H^X_{\alpha\beta}(0,0)=0$. We only compute $H^X_{\alpha\beta}(t,s)$ for $t>s$ because of the symmetry. It is straightforward to see that:

$$H^X_{\alpha\beta}(t+dt,s) = H^X_{\alpha\beta}(t,s)\left[1-\frac{dt}{\tau_\alpha}\right] + D^X_{\alpha\beta}(t,s)dt + o(dt),$$

with

$$D^X_{\alpha\beta}(t,s) = e^{-s/\tau_\alpha}\int_0^s e^{v/\tau_\alpha}\Delta^X_{\alpha\beta}(t,v)dv.$$

Hence computing $H^X_{\alpha\beta}(t+dt,s)$ knowing $H^X_{\alpha\beta}(t,s)$ amounts to computing $D_{\alpha\beta}(t,s)$. Fix $t\geq 0$. We have $D_{\alpha\beta}(t,0)=0$ and

$$D^X_{\alpha\beta}(t,s+ds) = D^X_{\alpha\beta}(t,s)(1-\frac{ds}{\tau_\alpha}) + \Delta^X_{\alpha\beta}(t,s)ds + o(ds).$$

This algorithm enables us to compute $H^X_{\alpha\beta}(t,s)$ for $t>s$. We deduce $H^X_{\alpha\beta}(t,s)$ for $t<s$ using the symmetry of this function. Finally, to get the values of $H^X_{\alpha\beta}(t,s)$ for $t=s$, we use the symmetry property of this function and get:

$$H^X_{\alpha\beta}(t+dt,t+dt) = H^X_{\alpha\beta}(t,t)\left[1-\frac{2dt}{\tau_\alpha}\right] + 2D^X_{\alpha\beta}(t,t)dt + o(dt).$$

These numerical schemes provide an efficient way for computing the mean and the covariance functions of the Gaussian process $F_1(X)$ (hence its probability distribution) knowing the law of the Gaussian process X. The algorithm used to compute the solution of the mean field equations for the general models GM1 and GMk is a straightforward generalization.

5.1.2 Analysis of the algorithm

Convergence rate As proved in theorem 3.8, given Z_0 a nondegenerate kP-dimensional Gaussian random variable and X a Gaussian process such that $X(0)=Z_0$, the sequences of means and covariance functions computed theoretically converge uniformly towards those of the unique fixed point of the map F_k. It is clear that our algorithm converges uniformly towards the real function it emulates. Hence for a finite N, the algorithm will converge uniformly towards the mean and covariance matrix of the process $F^N_k(X)$.

Denote by X_f the fixed point of F_k in $\mathcal{M}_1^+(C([t_0,T],\mathbb{R}^{kP}))$, of mean $\mu^{X_f}(t)$ and covariance matrix $C^{X_f}(t,s)$, and by $\widehat{F^N_k(X)}$ the numerical approximation of $F^N_k(X)$ computed using the algorithm previously described, whose mean is noted $\mu^{\widehat{F^N_k(X)}}(t)$ and whose covariance matrix is noted $C^{\widehat{F^N_k(X)}}(t,s)$. The uniform error between the simulated mean after N iterations with a time step dt and the fixed point's mean and covariance is the sum of the numerical error of the algorithm and the distance between the simulated process and the fixed point, is controlled by:

$$\|\mu^{\widehat{F^N_k(X)}} - \mu^{X_f}\|_\infty + \|C^{\widehat{F^N_k(X)}} - C^{X_f}\|_\infty = O(\ (N+T)\,dt + R_N(k_{\max})\ ) \qquad (35)$$

where $k_{\max} = \max(k, \tilde{k})$ and k and $\tilde{k}$) are the constants that appear in the proof of proposition 3.7 for the mean and covariance functions, and $R_N(x)$ is the exponential remainder, i.e. $R_N(x) = \sum_{n=N}^{\infty} x^n/n!$.

Indeed, we have:

$$\|\mu^{\widehat{F_k^N}(X)} - \mu^{X_f}\|_\infty \leq \|\mu^{\widehat{F_k^N}(X)} - \mu^{F_k^N(X)}\|_\infty + \|\mu^{F_k^N(X)} - \mu^{X_f}\|_\infty \tag{36}$$

The discretization algorithm used converges in $O(dt)$. Let us denote by C_1 the convergence constant, which depends on the sharpness of the function we approximate, which can be uniformly controlled over the iterations. Iterating the numerical algorithm has the effect of propagating the errors. Using these simple remarks we can bound the first term of the righthand side of (36), i.e. the approximation error at the Nth iteration:

$$\|\mu^{\widehat{F_k^N}(X)} - \mu^{F_k^N(X)}\|_\infty \leq C_1\, N\, dt$$

Because the sequence of means is a Cauchy sequence, we can also bound the second term of the righthand side of (36):

$$\begin{aligned}\|\mu^{F_k^N(X)} - \mu^{X_f}\|_\infty &\leq \sum_{n=N}^{\infty} \|\mu^{F_k^{n+1}(X)} - \mu^{F_k^n(X)}\|_\infty \\ &\leq \sum_{n=N}^{\infty} \frac{k^n}{n!} =: R_N(k)\end{aligned}$$

for some positive constant k introduced in the proof of proposition 3.7. The remainders sequence $(R_n(k))_{n\geq 0}$ converges fast towards 0 (an estimation of its convergence can be obtained using the fact that $\limsup_{k\to\infty}(1/k!)^{1/k} = 0$ by Stirling's formula).

Hence we have:

$$\|\mu^{\widehat{F_k^N}(X)} - \mu^{X_f}\|_\infty \leq C_1\, N\, dt + R_N(k) \tag{37}$$

For the covariance, the principle of the approximation is exactly the same:

$$\|C^{\widehat{F_k^N}(X)} - C^{X_f}\|_\infty \leq \|C^{\widehat{F_k^N}(X)} - C^{F_k^N(X)}\|_\infty + \|C^{F_k^N(X)} - C^{X_f}\|_\infty$$

The second term of the righthand side can be controlled using the same evaluation by $R_N(\tilde{k})$ where $\tilde{k}$ is the constant introduced in the proof of proposition 3.7, and the first term is controlled by the rate of convergence of the approximation of the double integral, which is bounded by $C_2(N + T)\, dt$ where C_2 depends on the parameters of the system and the discretization algorithm used.

Hence we have:

$$\|C^{\widehat{F_k^N}(X)} - C^{X_f}\|_\infty \leq C_2\, (N + T - t_0)\, dt + R_N(\tilde{k}) \tag{38}$$

The expressions (37) and (38) are the sum of two terms, one of which is increasing with N and T and decreasing with dt and the other one decreasing in N. If we want to obtain an estimation with an error bounded by some $\varepsilon > 0$, we can for instance fix N such that $\max(R_N(k), R_N(\tilde{k})) < \varepsilon/2$ and then fix the time step dt smaller than $\min(\, \varepsilon/(2C_1N), \varepsilon/(2C_2(N + T - t_0))\,)$.

Complexity The complexity of the algorithm depends on the complexity of the computations of the integrals. The algorithm described hence has the complexity $O(N(\frac{T}{dt})^2)$.

5.2 The importance of the covariance: Simple Model, one population.

As a first example and a benchmark for our numerical scheme we revisit the work of Sompolinsky and coworkers Sompolinsky et al (1988). These authors studied the case of the simple model with one population ($P = 1$), with the centered sigmoidal function $S(x) = \tanh(gx)$, centered connectivity weights $\bar{J} = 0$ of standard deviation $\sigma = 1$ and no input ($I = 0, \Lambda = 0$). Note therefore that there is no "noise" in the system, which therefore does not match the non degeneracy conditions of proposition 3.4 and of theorem 3.8 . This issue is discussed below. In this case, the mean equals 0 for all t. Nevertheless, the Gaussian process is non trivial as revealed by the study of the covariance $C(t, s)$.

5.2.1 Stationary solutions

Assuming that the solution of the mean field equation is a stationary solution with $C(t, s) \equiv C(t - s) = C(\tau)$, Sompolinsky and his collaborators found that the covariance obeyed a second order differential equation :

$$\frac{d^2C}{d\tau^2} = -\frac{\partial V_q}{\partial C}. \tag{39}$$

This form corresponds to the motion of a particle in a potential well and it is easy to draw the phase portrait of the corresponding dynamical system. However, there is a difficulty. The potential V_q depends on a parameter q which is in fact precisely the covariance at $\tau = 0$ ($q = C(0)$). In the stationary case, this covariance depends on the whole solution, and hence cannot be really considered as a parameter of the system. This is one of the main difficulties in this approach: mean field equations in the stationary regime are self-consistent.

Nevertheless, the study of the shape of V_q, considering q as a free parameter gives us some informations. Indeed, V_q has the following Taylor expansion (V_q is even because S is odd):

$$V_q(C) = \frac{\lambda}{2}C^2 + \frac{\gamma}{4}C^4 + O(C^6)$$

where $\lambda = (1 - g^2J^2\langle S'\rangle_q^2)$ and $\gamma = \frac{1}{6}J^2g^6\langle S^{(3)}\rangle_q^2)$, $\langle\phi\rangle_q$ being the average value of ϕ under the Gaussian distribution with mean zero and variance $q = C(0)$.

If $\lambda > 0$, i.e. when $g^2J^2\langle S'\rangle_q^2 < 1$, then the dynamical system (39) has a unique solution $C(t) = 0, \forall t \geq 0$. This corresponds to a stable fixed point (i.e. a deterministic trajectory, $\mu = 0$ with no fluctuations) for the neural network dynamics. On the other hand, if $g^2J^2\langle S'\rangle_q^2 \geq 1$ there is a homoclinic trajectory in (39) connecting the point $q = C^* > 0$ where V_q vanishes to the point $C = 0$. This solution is interpreted by the authors as a chaotic solution in the neural network. A stability analysis shows that this is the only stable[3] stationary solution Sompolinsky et al (1988).

[3]More precisely, this is the only minimum for the large deviation functional.

The equation for the homoclinic solution is easily found using energy conservation and the fact that $V_q(q) = 0$ and $\frac{dV_q}{dC}(q) = 0$. One finds:

$$u = \frac{dC}{dx} = -\sqrt{-V_q(C)}.$$

At the fourth order in the Taylor expansion of V_q this gives

$$C(\tau) = \frac{\sqrt{\frac{-2\lambda}{\gamma}}}{\cosh(\sqrt{-\frac{\lambda}{2}}\tau)}.$$

Though λ depends on q it can be used as a free parameter for interpolating the curve of $C(\tau)$ obtained from numerical data.

5.2.2 Numerical experiments

This case is a good benchmark for our numerical procedure since we know analytically the solutions we are searching for. We expect to find two regimes. In one case the correlation function is identically zero in the stationary regime, for sufficiently small g values or for a sufficiently small q (trivial case). The other case corresponds to a regime where $C(\tau) > 0$ and $C(\tau) \to 0$ has $\tau \to +\infty$ ("chaotic" case). This regime requires that g be sufficiently large and that q be large too. We took $\tau_\alpha = 0.25, \sigma_{\alpha\alpha} = 1$. For these values, the change in dynamics predicted by Sompolinsky and collaborators is $g_c = 4$.

In sections 3 and 4 we have introduced the assumption of non-degeneracy of the noise, in order to ensure that the mean field process was non degenerate. However, in the present example, there is no external noise in the evolution, so we can observe the effects of relaxing this hypothesis in a situation where the results of proposition 3.4 and of theorem 3.8 cannot be applied. First, we observed numerically that, without external noise, the process could become degenerate (namely some eigenvalues of the covariance matrix $C_\alpha(t, s)$ become very small and even vanish.). This has also an incidence on the convergence of the method which presents numerical instabilities, though the iterations leads to a curve which is well fitted by the theoretical results of Sompolinsky et al. (see Fig. 3) . The instability essentially disappears if one adds a small noise. But, note that in this case, the solution does not match with Sompolinsky et al. theoretical calculation (see Fig. 3).

Modulo this remark, we have first considered the trivial case corresponding to small g values. We took $g = 0.5$ and $T = 5$. We choose as initial process the stationary Ornstein-Uhlenbeck process corresponding to the uncoupled system with $\Lambda = 0.1$. We drew $\mu_\alpha(0)$ randomly from the uniform distribution in $[-1, 1]$ and $v_\alpha(0)$ randomly from the uniform distribution in $[0, 1]$.

Starting from this initial stationary process, we iterated the function F_1. Then, during the iterations, we set $s_\alpha = 0$ in order to match the conditions imposed by Sompolinsky and coworkers. We observe that the method converges towards the expected solution: the mean function converges to zero, while the variance $v(t)$ decreases exponentially fast in time towards a constant value corresponding to the stationary regime. This asymptotic value decreases between two consecutive iterations, which is consistent with the theoretical expectation that $v(t) = 0$ in the stationary regime of

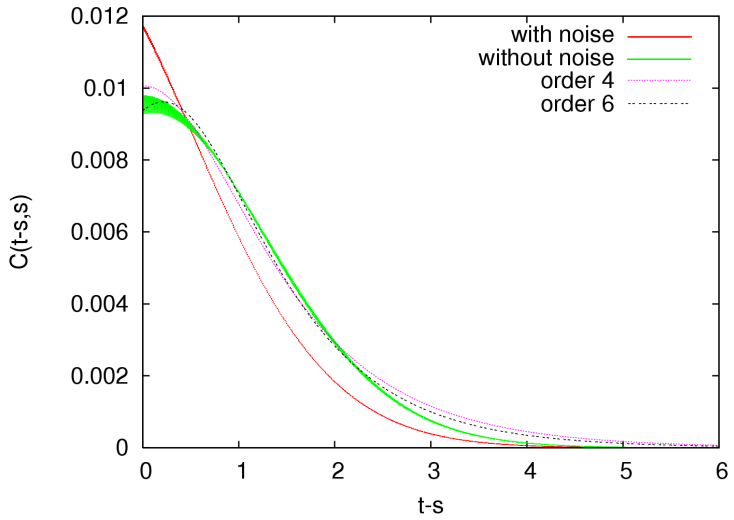


Figure 3: Numerical solution of the mean field equation after 14 iterations in the chaotic case ($g = 5$). We clearly see the numerical instabilities in the no-noise case, which do not exist in the low-noise case.

the trivial case. Finally, we observe that the covariance $C(t-s,s)$ stabilizes to a curve that does not depend on s and the stationary value (large $t-s$) converges to zero.

We applied the same procedure for $g = 5$ corresponding to the "chaotic" regime. The behavior was the same for $\mu(t)$ but was quite different for the covariance function $C(t,s)$. Indeed, while in the first case the stationary value of $v(t)$ tends to zero with the number of iterations, in the chaotic case it stabilizes to a finite value. In the same way, the covariance $C(t-s,s)$ stabilizes to a curve that does not depend on s. The shape of this curve can be extrapolated thanks to Sompolinsky et al. results. We observe a very good agreement with the theoretical predictions with a fit $f_4(x) = \frac{a}{\cosh(b(x-\delta))}$, corresponding to the fourth expansion of V_q. Using a 6-th order expansion of $V_q(x) = \frac{a}{2}x^2 + \frac{b}{4}x^4 + \frac{c}{6}x^2$ gives a fit $f_6(x) = \frac{\rho}{\cosh(\lambda(x-\delta))}\frac{1}{\sqrt{1+K^2-\frac{1}{\cosh^2(\lambda(x-\delta))}}}$, where ρ, K, λ are explicit functions of a, b, c, we obtain a slightly better approximation.

5.3 Mean field equations for two populations with a negative feedback loop.

Let us now present a case where the fluctuations of the Gaussian field act on the dynamics of $\mu_\alpha(t)$ in a non trivial way, with a behavior strongly departing from the naive mean field picture. We consider two interacting populations where the connectivity weights are Gaussian random variables $J_{\alpha\beta} \equiv \mathcal{N}(\bar{J}_{\alpha\beta}, \sigma_{\alpha\beta} = 1)$ for $(\alpha,\beta) \in \{1,2\}^2$. We set $S_{\alpha\beta}(x) = \tanh(gx)$ and $I_\alpha = 0, s_\alpha = 0, \alpha = 1,2$.

5.3.1 Theoretical framework.

The dynamic mean field equation for $\mu_\alpha(t)$ is given, in differential form, by:

$$\frac{d\mu_\alpha}{dt} = -\frac{\mu_\alpha}{\tau_\alpha} + \sum_{\beta=1}^{2} \bar{J}_{\alpha\beta} \int_{-\infty}^{\infty} S\left(\sqrt{v_\beta(t)}x + \mu_\beta(t)\right) Dx,\ \alpha = 1, 2.$$

Let us denote by $G_\alpha(\mu, v(t))$ the function in the righthand side of the equality. Since S is odd, $\int_{-\infty}^{\infty} S(\sqrt{v_\beta(t)}x)Dx = 0$. Therefore, we have $G_\alpha(0, v(t)) = 0$ whatever $v(t)$, and hence the point $\mu_1 = 0, \mu_2 = 0$ is always a fixed point of this equation.

Let us study the stability of this fixed point. To this purpose, we compute the partial derivatives of $G_\alpha(\mu, v(t))$ with respect to μ_β for $(\alpha, \beta) \in \{1, 2\}^2$. We have:

$$\frac{\partial G_\alpha}{\partial \mu_\beta}(\mu, v(t)) = -\frac{\delta_{\alpha\beta}}{\tau_\alpha} + g\bar{J}_{\alpha\beta} \int_{-\infty}^{\infty} \left(1 - \tanh^2\left(\sqrt{v_\beta(t)}x + \mu_\beta(t)\right)\right) Dx,$$

and hence at the point $\mu_1 = 0, \mu_2 = 0$, these derivatives read:

$$\frac{\partial G_\alpha}{\partial \mu_\beta}(0, v(t)) = -\frac{\delta_{\alpha\beta}}{\tau_\alpha} + g\bar{J}_{\alpha\beta} h(v_\beta(t)),$$

where $h(v_\beta(t)) = 1 - \int_{-\infty}^{\infty} \tanh^2(\sqrt{v_\beta(t)}x)Dx$.

In the case $v_\alpha(0) = 0, J = 0, s_\alpha = 0$, implying $v_\alpha(t) = 0, t \geq 0$, the equation for μ_α reduces to:

$$\frac{d\mu_\alpha}{dt} = -\frac{\mu_\alpha}{\tau_\alpha} + \sum_{\beta=1}^{2} \bar{J}_{\alpha\beta} S(\mu_\beta(t))$$

which is the standard Amari-Cohen-Grossberg-Hopfield system. This corresponds to the naive mean field approach where Gaussian fluctuations are neglected. In this case the stability of the fixed point $\mu = 0$ is given by the sign of the largest eigenvalue of the Jacobian matrix of the system that reads:

$$\begin{pmatrix} -\frac{1}{\tau_1} & 0 \\ 0 & -\frac{1}{\tau_2} \end{pmatrix} + g \begin{pmatrix} \bar{J}_{11} & \bar{J}_{12} \\ \bar{J}_{21} & \bar{J}_{22} \end{pmatrix}.$$

For the sake of simplicity we assume that the two time constants τ_α are equal and we denote this value τ. The eigenvalues are in this case $-\frac{1}{\tau} + g\lambda$, where λ are the eigenvalues of $\bar{\mathcal{J}}$ and have the form:

$$\lambda_{1,2} = \frac{\bar{J}_{11} + \bar{J}_{22} \pm \sqrt{(\bar{J}_{11} - \bar{J}_{22})^2 + 4\bar{J}_{12}\bar{J}_{21}}}{2}.$$

Hence, they are complex whenever $\bar{J}_{12}\bar{J}_{21} < -(\bar{J}_{11} - \bar{J}_{22})^2/4$, corresponding to a negative feedback loop between population 1 and 2. Moreover, they have a real part only if $\bar{J}_{11} + \bar{J}_{22}$ is non zero (self interaction).

This opens up the possibility to have an instability of the fixed point ($\mu = 0$) leading to a regime where the average value of the membrane potential oscillates. This occurs if $\bar{J}_{11} + \bar{J}_{22} > 0$ and if g is larger than:

$$g_c = \frac{2}{\tau(\bar{J}_{11} + \bar{J}_{22})}.$$

The corresponding bifurcation is a Hopf bifurcation.

The situation is different if one takes into account the fluctuations of the Gaussian field. Indeed, in this case the stability of the fixed point $\mu = 0$ depends on $v(t)$. More precisely, the real and imaginary part of the eigenvalues of $DG(0, v(t))$ depend on $v(t)$. Therefore, the variations of $v(t)$ act on the stability and oscillations period of $v(t)$. Though the evolution of $\mu(t), v(t)$ are coupled we cannot consider this evolution as a coupled dynamical system, since $v(t) = C(t, t)$ is determined by the mean field equation for $C(t, s)$ which cannot be written as an ordinary differential equation. Note that we cannot assume stationarity here, as in the previous case, since $\mu(t)$ depends on time for sufficiently large g. This opens up the possibility of having complex dynamical regimes when g is large.

5.3.2 Numerical experiments

We have considered the case $\bar{J}_{11} = \bar{J}_{22} = 5$,$\tau = 0.1$ giving a Hopf bifurcation for $g_c = 2$ when $J = 0$ (fig. 4). The trajectory of $\mu_1(t)$ and $v_1(t)$ is represented in Figure 4 in the case $g = 3$. When $J = 0$, $\mu_1(t)$ presents regular oscillations (with non linear effects since $g = 3$ is larger than the critical value for the Hopf bifurcation, $g_c = 2$). In this case, the solution $v_1(t) = 0$ is stable as seen on the figure. When $J \neq 0$ the Gaussian field has (small) fluctuations which nevertheless strongly interact with the dynamics of $\mu_1(t)$, leading to a regime where $\mu_1(t)$ and $v_1(t)$ oscillate periodically

6 Conclusion

The problem of bridging scales is overwhelming in general when studying complex systems and in particular in neuroscience. After many others we look at this difficult problem from the theoretical and numerical viewpoints, hoping to get closer to its solution from relatively simple and physically/biologically plausible first principles and assumptions. One of our motivations is to better understand such phenomenological neural mass models as that of Jansen and Rit Jansen and Rit (1995).

We consider several populations of neurons and start from a microscopic, i.e. individual, description of the dynamics of the membrane potential of each neuron that contains four terms. The first one controls the intrinsic dynamics of the neuron. It is linear in this article but this assumption is not essential and could probably be safely removed if necessary. The second term is a stochastic input current, correlated or uncorrelated. The third one is a deterministic input current, and the fourth one describes the interaction between the neurons through random connectivity coefficients that weigh the contributions of other neurons through a set of functions that are applied to their membranes potentials. The only hypothesis on these functions is that they are smooth and bounded. The obvious

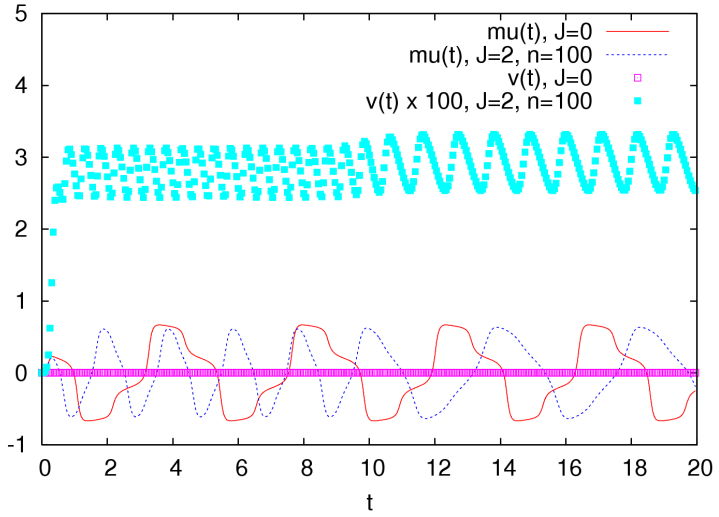


Figure 4: Evolution of the mean $\mu_1(t)$ and variance $v_1(t)$ for the mean field of population 1, for $J = 0$ and $J = 2$, over a time window $[0, 20]$. n is the number of iterations of F_1 defined in section 3. This corresponds to a number of iterations for which the method has essentially converged (up to some precision). Note that $v_1(t)$ has been magnified by a factor of 100. Though Gaussian fluctuations are small, they have a strong influence on $\mu_1(t)$.

choice of sigmoids is motivated by standard rate models ideas. Another appealing choice is a smooth approximation to a Dirac delta function thereby opening a window on the world of spiking neurons.

We then derive the mean field equations and provide a constructive and new proof, under some mild assumptions, of the existence and uniqueness of a solution of these equations over finite and infinite time intervals. The key idea is to look at this mean field description as a *global problem* on the probability distribution of the membranes potentials, unlike previous studies. Our proof provides an efficient way of computing this solution and our numerical experiments show a good agreement with previous studies.

In the case where the nonlinearities are chosen to be sigmoidal our results shed a new and fascinating light on existing neural mass models. Indeed these appear as approximations of the mean field equations where the intricate but fundamental coupling between the time variations of the mean membrane potentials and their fluctuations, as represented by the covariance functions, is neglected. This article is just a small step toward answering from the theoretical and numerical standpoints the questions raised by this coupling but we are convinced that a host of interesting results can be found there.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Aknowledgement

This research was partly supported by funding of the European Union under the grant no. 15879 (FACETS) and the Fondation d'Entreprise EADS.

A Identification of the mean field equations

Ben-Arous and Guionnet studied from a mathematical point of view the problem of finding a mean-field description of large networks of spin glasses. They obtained using different methods of stochastic analysis a weak limit of the law of a given spin and proved their independence.

Our equations do not directly fit in their study: indeed, the spin intrinsic dynamics is nonlinear while the interaction is linear, and everything in done in dimension one. Nevertheless, their proof extends to our case which is somehow more simple. For instance in the case of the Simple Model with one population, we can readily adapt their proof in our case. More precisely, let $P = 1$, the equation of the network reads:

$$\tau dV_t^j = (-V_t^j + \sum_{i=1}^{N} J_{ij} S(V_t^i))\, dt + \sigma dW_t^j$$

In this case, we define for $X \in \mathcal{M}_1^+(C([t_0, T], \mathbb{R})$ the effective interaction term (U_t^X) which is the effective interaction process defined in 2.1, i.e. the Gaussian process of mean $\bar{J}_{\alpha\beta}\mathbb{E}\left[S(X_t)\right]$ and of covariance: $\mathrm{Cov}\left(U_t^X,\, U_s^X\right) =: \sigma_{\alpha\beta}^2 \mathbb{E}\left[S(X_t)S(X_s)\right]$.

Let us note $\mathcal{P}$ the law of the membrane potential when there is no interaction (it is an Ornstein-Ulhenbeck process), and the empirical measure $\hat{V}^N = \frac{1}{N}\sum_{i=1}^N \delta_{V^i}$. We can prove that under the probability distribution averaged over the connectivities, see below, the empirical measure satisfies a large deviation principle with good rate function H defined as in (Guionnet, 1997). Using this large deviation result, we can prove annealed and quenched tightness of the empirical measure, and finally its convergence towards the unique process where the good rate function H achieves its unique minimum, which is defined by the property of having a density with respect to $\mathcal{P}$ and whose density satisfies the implicit equation:

$$Q \ll \mathcal{P} \qquad \frac{dQ}{d\mathcal{P}} = \mathcal{E}\left[\exp\left\{\int_0^T U_t^Q dW_t - \frac{1}{2}\int_0^T (U_t^Q)^2\, dt\right\}\right] \tag{40}$$

where $\mathcal{E}$ denotes the expectation over the effective interaction process U^Q.

We can also prove following the steps of Ben-Arous and Guionnet in (Ben-Arous and Guionnet, 1997) that there exists a unique solution to this equation, and that this solution satisfies the nonlinear nonmarkovian stochastic differential equation:

$$\begin{cases} \tau dV_t = -V_t\,dt + dB_t \\ dB_t = dW_t + \int_0^t dB_s \mathcal{E}\left[U_s^Q U_t^Q \frac{\exp\{-\frac{1}{2}\int_0^t (U_u^Q)^2 du\}}{\mathcal{E}[\exp\{-\frac{1}{2}\int_0^t (U_u^Q)^2 du\}]}\right] \\ \text{Law of } (V) = Q, \ \text{ law of } (V_0) = Z_0 \end{cases} \tag{41}$$

which can also be written as our mean field equation, averaged on the connectivities (see (Ben-Arous and Guionnet, 1995)). More precisely, let L^V be the law of the solution of the equation:

$$\begin{cases} \tau dV_t = -V_t dt + dW_t + U_t^V dt \\ \text{Law of } V_0 = Z_0 \end{cases},$$

which is exactly equation (13). They prove that V satisfies the nonlinear equation:

$$V \overset{\mathcal{L}}{=} \mathcal{E}(L^V)$$

This result is likely extendable to the multi population case but the corresponding mathematical developments are out of the scope of this paper.

B The resolvent

In this appendix we introduce and give some useful properties of the resolvent Φ_L of a homogeneous differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mathbf{L}(t)x(t) \quad x(t_0) = x_0 \in \mathbb{R}^P, \tag{42}$$

where $\mathbf{L} : [t_0, T] \to \mathcal{M}_{P\times P}$ (or $(-\infty, T] \to \mathcal{M}_{P\times P}$) is C^0.

Definition B.1. The resolvent of (42) is defined as the unique solution of the linear equation:

$$\begin{cases} \frac{\mathrm{d}\Phi_L(t,t_0)}{\mathrm{d}t} & = \mathbf{L}(t)\Phi_L(t,t_0) \\ \Phi_L(t_0,t_0) & = \mathrm{Id}_P \end{cases} \tag{43}$$

where Id_P is the $P \times P$ identity matrix.

Proposition B.1. The resolvent satisfies the following properties:

(i). $\Phi_L(t+s,t_0) = \Phi_L(t+s,t) \cdot \Phi_L(t,t_0)$

(ii). $\Phi_L(t,t_0)$ is invertible of inverse $\Phi_L(t_0,t)$ which satisfies:

$$\begin{cases} \frac{\mathrm{d}\Phi_L(t_0,t)}{\mathrm{d}t} & = -\Phi_L(t_0,t)\mathbf{L}(t) \\ \Phi_L(t_0,t_0) & = \mathrm{Id}_{P\times P} \end{cases} \tag{44}$$

(iii). Let $\|\ \|$ be a norm on $\mathcal{M}_{P\times P}$ and assume that $\|\mathbf{L}(t)\| \leq k_L$ on $[t_0, T]$. Then we have:

$$\|\Phi_L(t, t_0)\| \leq e^{k_L|t-t_0|} \quad \forall t \in [t_0, T] \tag{45}$$

Similarly, if $\left\|\mathbf{L}^T(t)\right\| \leq k_{L^T}$ on $[t_0, T]$ we have:

$$\left\|\Phi_L^T(t, t_0)\right\| \leq e^{k_{L^T}|t-t_0|} \quad \forall t \in [t_0, T] \tag{46}$$

(iv). We have

$$\det\Phi_L(t, t_0) = \exp \int_{t_0}^{t} \mathrm{Tr}\mathbf{L}(s)\, ds$$

Proof. The properties (i) and (ii) are directly linked with the property of group of the flow of a reversible ODE. (iii) is an application of Gronwald's lemma. (iv) is obtained by a first order Taylor series expansion. □

We also need in the article a lower bound on $\|\Phi_L(t, t_0)\|$ for all $t \in [t_0, T]$ in the general case where $\mathbf{L}$ is not constant. This can be achieved for example using Floquet's theory. Consider the interval $[t_0, 2T - t_0]$ and define the continuous periodic function $\widetilde{\mathbf{L}}(t)$ of period $2(T - t_0)$ defined by

$$\widetilde{\mathbf{L}}(t) = \begin{cases} \mathbf{L}(t) & t_0 \leq t \leq T \\ \mathbf{L}(2T - t) & T \leq t \leq 2T - t_0 \end{cases}$$

The corresponding resolvent $\Phi_{\tilde{L}}(t, t_0)$ is equal to $\Phi_L(t, t_0)$ for $t_0 \leq t \leq T$. $\Phi_{\tilde{L}}(2T - t_0, t_0)$ is invertible and hence there exists $a \in \mathbb{R}$ such that

$$e^{2a(T-t_0)} < |\lambda|$$

for all eigenvalues λ of $\Phi_{\tilde{L}}(2T - t_0, t_0)$. One of Floquet's theorems states that there exists a norm on $\mathbb{R}^P$ and $\gamma > 0$ such that

$$\gamma e^{a(t-t_0)} < \|\Phi_{\tilde{L}}(t, t_0)\| \quad t \geq t_0,$$

and in particular

$$\gamma e^{a(t-t_0)} < \|\Phi_L(t, t_0)\| \quad t_0 \leq t \leq T \tag{47}$$

Theorem B.2 (Solution of an inhomogeneous linear SDE)**.** *The solution of the inhomogeneous linear Stochastic Differential Equation:*

$$\begin{cases} dX_t &= (\mathbf{L}(t)X(t) + \mathbf{I}(t))\, dt + \mathbf{\Lambda}(s)d\mathbf{W}_s \\ X_{t_0} &= X_0 \end{cases} \tag{48}$$

can be written using the resolvent:

$$X_t = \Phi_L(t, t_0)X_0 + \int_{t_0}^{t} \Phi_L(t, s)\mathbf{I}(s)\, ds + \int_{t_0}^{t} \Phi_L(s, t)\mathbf{\Lambda}(s)d\mathbf{W}_s \tag{49}$$

Proof. Pathwise (strong) uniqueness of solution directly comes from the results on the SDE with Lipschitz coefficients (see e.g. (Karatzas and Shreve, 1991, Theorem 2.5 of Chapter 5)). It is clear that $X_{t_0} = X_0$. We use Itô's formula for the product of two stochastic processes to prove that the process (49) is solution of equation (48):

$$\begin{aligned}
dX_t &= \Big(\mathbf{L}(t)\Phi_L(t,t_0)X_0 + \Phi_L(t,t)\mathbf{I}(t) + \int_{t_0}^t \mathbf{L}(t)\Phi_L(t,s)\mathbf{I}(s)\,ds\Big)\,dt \\
&\quad + \Phi_L(t,t)\mathbf{\Lambda}(t)d\mathbf{W}_t + \int_{t_0}^t \mathbf{L}(t)\Phi_L(s,t)\mathbf{\Lambda}(s)d\mathbf{W}_s\,dt \\
&= \Big(\mathbf{L}(t)\Big[\Phi_L(t,t_0)X_0 + \int_{t_0}^t \Phi_L(s,t)I(s)\,ds + \int_{t_0}^t \Phi_L(s,t)\mathbf{\Lambda}(s)d\mathbf{W}_s\Big] + \mathbf{I}(t)\Big)\,dt \\
&\quad + \mathbf{\Lambda}(t)d\mathbf{W}_t \\
&= \left(\mathbf{L}(t)X(t) + \mathbf{I}(t)\right)dt + \mathbf{\Lambda}(t)d\mathbf{W}_t
\end{aligned}$$

Hence the theorem is proved. □

C Matrix norms

In this section we recall some definitions on matrix and vector norms. Let $\mathcal{M}_{n\times n}$ be the set of $n \times n$ real matrices. It is a vector space of dimension n^2 and the usual L^p norms $1 \leq p \leq \infty$ can be defined. Given $\mathbf{L} \in \mathcal{M}_{n\times n}$, we note $\|\mathbf{L}\|_p^v$ the corresponding norm. Given a vector norm, noted $\|\ \|$, on $\mathbb{R}^n$ the induced norm, noted $\|\ \|$, on $\mathcal{M}_{n\times n}$ is defined as

$$\|\mathbf{L}\| = \sup_{x\in\mathbb{R}^n,\,\|x\|\leq 1} \frac{\|\mathbf{L}x\|}{\|x\|}$$

Since $\mathcal{M}_{n\times n}$ is finite dimensional all norms are equivalent. In this article we use the following norms

(i). $\|\mathbf{L}\|_\infty = \max_i \sum_{j=1}^n |L_{ij}|$.

(ii). $\|\mathbf{L}\|_\infty^v = \max_{i,\,j} |L_{ij}|$

(iii). $\|\mathbf{L}\|_2 = \sup_{x\in\mathbb{R}^n,\,\|x\|_2\leq 1} \frac{\|\mathbf{L}x\|_2}{\|x\|_2}$. This so-called *spectral* norm is equal to the square root of the largest singular value of $\mathbf{L}$ which is the largest eigenvalue of the positive matrix $\mathbf{L}^T\mathbf{L}$. If $\mathbf{L}$ is positive definite this is its largest eigenvalue which is also called its spectral radius, noted $\rho(\mathbf{L})$.

D Important quantities

Table 1 summarizes some notations which are introduced in the article and used in several places.

Constant	Defined in
μ	lemma 3.2 equation (21)
$\sigma_{\max}$	lemma 3.2
$\sigma_{\min}$	lemma 3.2
k_0	lemma 3.5
K	proof of lemma 3.6
k_C	proposition 3.7 equation (27)
λ_L	equation (30)

Table 1: Some important quantities defined in the article.

References

Abbott L, Van Vreeswijk C (1993) Asynchronous states in networks of pulse-coupled neuron. Phys Rev 48:1483–1490

Amari S (1972) Characteristics of random nets of analog neuron-like elements. Syst Man Cybernet SMC-2

Amari SI, Yoshida K, Kanatani KI (1977) A mathematical foundation for statistical neurodynamics. Siam J Appl Math 33(1):95–126

Ben-Arous G, Guionnet A (1995) Large deviations for Langevin spin glass dynamics. Probability Theory and Related Fields 102(4):455–509

Ben-Arous G, Guionnet A (1997) Symmetric Langevin Spin Glass Dynamics. The Annals of Probability 25(3):1367–1422

Bogachev V (1998) Gaussian Measures. American Mathematical Society

Braitenberg V, Schüz A (1998) Cortex: Statistics and Geometry of Neuronal Connectivity, 2nd edn. Springer

Brunel N, Hakim V (1999) Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. Neural Computation 11:1621–1671

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

Crisanti A, Sompolinsky H (1987a) Dynamics of spin systems with randomly asymmetric bonds: Langevin dynamics and a spherical model. Physical Review A 36(10):4922–4939

Crisanti A, Sompolinsky H (1987b) Dynamics of spin systems with randomly asymmetric bounds: Ising spins and glauber dynamics. Phys Review A 37(12):4865

Crisanti A, Sommers H, Sompolinsky H (1990) chaos in neural networks : chaotic solutions

Dayan P, Abbott LF (2001) Theoretical Neuroscience : Computational and Mathematical Modeling of Neural Systems. MIT Press

Ermentrout B (1998) Neural networks as spatio-temporal pattern-forming systems. Reports on Progress in Physics 61:353–430

Freeman W (1975) Mass action in the nervous system. Academic Press, New York

Gerstner W (1995) Time structure of the activity in neural network models. Physical Review E 51(1):738–758

Gerstner W, Kistler WM (2002) Mathematical formulations of hebbian learning. Biological Cybernetics 87:404–415

Grimbert F, Faugeras O (2006) Bifurcation analysis of Jansen's neural mass model. Neural Computation 18(12):3052–3068

Guionnet A (1997) Averaged and quenched propagation of chaos for spin glass dynamics. Probability Theory and Related Fields 109(2):183–215

Hopfield JJ (1984) Neurons with graded response have collective computational properties like those of two-state neurons. Proceedings of the National Academy of Sciences, USA 81(10):3088–3092

Houghton A, Jain S, Young AP (1983) Role of initial conditions in the mean-field theory of spin-glass dynamics. Phys Rev B 28(5):2630–2637

Jansen BH, Rit VG (1995) Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biological Cybernetics 73:357–366

Karatzas I, Shreve SE (1991) Brownian motion and stochastic calculus, Graduate Texts in Mathematics, vol 113, 2nd edn. Springer-Verlag, New York

Mattia M, Del Giudice P (2002) Population dynamics of interacting spiking neurons. Physical Review E 66(5):51,917

Molgedey L, Schuchardt J, Schuster H (1992) Supressing chaos in neural networks by noise. Physical Review Letters 69(26):3717–3719

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

van Rotterdam A, Lopes da Silva F, van den Ende J, Viergever M, Hermans A (1982) A model of the spatial-temporal characteristics of the alpha rhythm. Bulletin of Mathematical Biology 44(2):283–305

Samuelides M, Cessac B (2007) Random recurrent neural networks. European Physical Journal - Special Topics 142:7–88

Sompolinsky H, Zippelius A (1982) Relaxational dynamics of the Edwards-Anderson model and the mean-field theory of spin-glasses. Physical Review B 25(11):6860–6875

Sompolinsky H, Crisanti A, Sommers H (1988) Chaos in Random Neural Networks. Physical Review Letters 61(3):259–262

Treves A (1993) Mean-field analysis of neuronal spike dynamics. Network: Computation in Neural Systems 4(3):259–284



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399