arXiv:1601.00985v1 [math.PR] 5 Jan 2016

Large deviations of particle systems in random interaction

Tanguy Cabana   Jonathan Touboul

January 7, 2016

Abstract

We investigate the thermodynamic limit of a class of particle systems in random interaction that encompasses coupled oscillators systems and neuronal networks. In these systems, the interactions depend asymmetrically on the state of both particle and its amplitude is scaled by a Gaussian random coefficient whose variance decays as the inverse of the network size. We show that the empirical measure satisfies a large-deviation principle with good rate function achieving its minimum at a unique probability measure, implying convergence of the empirical measure and propagation of chaos. The limit is characterized through a complex non Markovian implicit equation in which the network interaction term is replaced by a Gaussian field depending on the state of the particle.

Work in progress - Preliminary version of January 7, 2016

Contents

1 Introduction

Interacting particle systems in random environments are ubiquitous in the theory of complex systems. They are useful to model a broad range of phenomena, from neural networks of the brain [26] to communication networks [6], internet traffic [17], disordered physical systems [3] and economics and social science [24]. A particularly important model in life science is the Kuramoto model of coupled oscillators [20] which is one of the seldom models that are completely solvable. All these models have in common to be generally described by a variable $(X_t^{i,N})_{i=1\cdots N} \in \mathbb{R}^N$ governing the state of each particle, and that satisfies a stochastic differential equation of type:

$$dX_t^{i,N} = \left(f(r_i, t, X_t^{i,N}) + \sum_{j=1}^{N} J_{ij} b(X_t^{i,N}, X_t^{j,N}) \right) dt + \lambda dW_t^i. \tag{1}$$

In that equation, each particle has its own intrinsic dynamics governed by the map $f(r_i,\cdot,\cdot)$, where r_i accounts for the heterogeneous nature of the different particles; it is also subject to stochastic fluctuations governed a Brownian motion $(W_t^i)_{t\geq 0}$ independent of the fluctuations of other particles. The summation term describes the interactions: it is given by the product of a typical interaction term $b(X_t^{i,N}, X_t^{j,N})$ modulated by a coefficient $J_{ij}\in\mathbb{R}$ governing the amplitude of the interaction of particle j onto i.

The heterogeneities of the intrinsic dynamics are classically taken into account by considering that $(r_i)_{i=1\cdots N}$ are identically distributed (i.i.d.) random variables $r_i\in D, 1\leq i\leq N$, where $D\subseteq\mathbb{R}^d$ for some $d\in\mathbb{N}^*$ with distribution $\pi\in\mathcal{M}_1^+(D)$ absolutely continuous with respect to Lebesgue's measure. Similarly, the heterogeneous interaction amplitude are i.i.d. random variables. In order for the interaction term to remain non-trivial (i.e., neither diverge nor disappear) in the thermodynamic limit, their law depends crucially on the network size N. It is thus natural to consider that their mean is equivalent to $\frac{\bar{J}}{N}$ with $\bar{J}\in\mathbb{R}$. A canonical situation that we refer to as the mean-field interactions model corresponds to the case where the variance is negligible compared to $\frac{1}{N}$ (typically, this covers the case where these coefficients are deterministic). In that case, under sufficient regularity conditions on the intrinsic dynamics and interaction maps, it is well known [30, 13, 22, 21] that the system enjoys the propagation of chaos property and converges towards a McKean-Vlasov equation

$$d\bar{X}_t(r) = \left(f(r,t,\bar{X}_t(r)) + \bar{J}\int b(\bar{X}_t(r),z)p(t,r',dz)d\pi(r')\right)dt + \lambda dW_t, \tag{2}$$

where p is the law of $\bar{X}$. We note that in this limit, the interaction term has become implicit and deterministic, and the possible randomness of microscopic interactions do not affect in any way the behavior of the macroscopic system.

There is only one regime preserving a nontrivial contribution of microscopic disorder in the thermodynamic limit: this occurs when the variance of the interaction amplitudes is equivalent to σ^2/N. This regime is particularly rich, nonstandard, and still not fully understood. In this scaling, it was shown in different situations that the level of disorder has an important impact on the macroscopic behavior: in spin glass systems, this parameter governs the glassy transition [25], in randomly connected neural networks, this parameter governs a transition from trivial states to chaotic dynamics [26], and in the Kuramoto model, a very singular and somewhat controversial transition occurs as disorder increase [10, 11, 27, 28]. Mathematically, important advances were achieved in the characterization of the thermodynamics limit of Langevin spin glass systems by Ben-Arous, Dembo and Guionnet [3, 18, 1, 4]. These works are fundamental in the field in that they introduce a general methodology to characterize systems of bounded spins with linear interactions only depending on the state of the other particles (i.e., $b(x,y)=y$). Their results prove large deviations properties on the empirical measure, propagation of chaos and convergence towards a non-Markov implicit equation, and in the spherical spin glass case, aging [1]. The same technique was used in neuroscience in order to understand the dynamics rate models in which the interactions also only depend on pre-synaptic cells via a sigmoid transform, $b(x,y)=S(y)$, in discrete-time systems [23, 16], and then developed to incorporate multiple populations, spatial extension and delays [7, 8]. These results would apply to system (1) with $b(x,y)\equiv B(y)$ sufficiently regular, and would prove a convergence towards an equation of type:

$$d\bar{X}_t(r) = \left(f(r,t,\bar{X}_t(r)) + U_t^{\bar{X}}\right)dt + \lambda dW_t, \tag{3}$$

where $U_t^{\bar{X}}$ would be a Gaussian process with mean and covariance:

$$\begin{cases}\mathbb{E}[U_t^{\bar{X}}] = \bar{J}\int_D \mathbb{E}[B(\bar{X}_t(r'))]d\pi(r')\\ \mathrm{Cov}(U_t^{\bar{X}},U_s^{\bar{X}}) = \sigma^2\int_D \mathbb{E}[B(\bar{X}_t(r'))B(\bar{X}_s(r'))d\pi(r').\end{cases}$$

In applications, it is important to consider cases in which $b(x,y)$ do depend on x. This is the case in most accurate biophysical descriptions of neural networks [19, 14], of interacting oscillators [20], of swarming models [9] and gases [5]. However, even from the physical viewpoint, it is not clear how previous mathematical results extend to maps $b(x,y)$ depending on x. Indeed, under Boltzmann's molecular chaos ansatz - particles are iid and independent of the interaction amplitudes - and assuming $b(x, y) \equiv B(y)$, the terms $\sum_j J_{ij} b(X_t^{i,N}, X_t^{j,N})$ are likely to converge towards $U_t^{\bar{X}}$ by the virtue of a functional central limit theorem. But this is no longer true in the general case since the terms in the sum would no more be independent. Moreover, is not immediately clear what could be a possible limit for that system.

We undertake here the mathematical analysis of the limits of equation (1) in the general case. In the case of Gaussian interaction amplitudes, we will show that the averaged empirical measure satisfies a weak large-deviation principle with a good rate function achieving its minimum at a unique probability measure for sufficiently small times. This implies both convergence and propagation of chaos of the system. While a similar short-time restriction for spin glass systems [3] was relaxed to the price of a weak large-deviation principle [18], the dependence in x of the interaction terms seems to prevent such extensions.

The paper is organized as follow. We introduce our mathematical setting in Section 2 and state our results. Section 3 provides the proof of our weak Large Deviations Principle that relies on the identification of the good rate function, an upper-bound result for compact sets, as well as a tightness result. In Section 4, we demonstrate that the good rate function admits a unique minimum, that it is also the unique solution of (3), and proves the convergence of the empirical measure toward it.

2 Mathematical setting and main results

The interacting particle system (1) is a diffusion in random environment, and as such involves two probability spaces:

- The intrinsic dynamics of the particles as well as their interaction amplitudes are random variables on a complete probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \mathcal{P})$. These heterogeneities are taken into account through the independent random variables $(r_i)_{i\in\mathbb{N}}$ and $(J_{ij})_{i,j} \in \mathbb{N}^2$. They constitute the random environment of the dynamic and are frozen in time. Their realization do not depend on the evolution of the system. We will denote $\mathcal{E}$ the expectation under $\mathcal{P}$, and by $\mathcal{E}_J$ and $\mathcal{P}_J$ the expectation and probability over the variables J_{ij} only.
- The particles are driven by $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ independent standard Brownian motions (W_t^i).

The resulting dynamics of the particles thus depends both on the random environment and on the Brownian motions. We assume that the parameters of the dynamics of the network equations enjoy the following regularity assumptions:

1. The function f is K_f-Lipschitz-continuous in its two variables.
2. The function b is bounded and K_b-Lipschitz-continuous in both variables. We note $\|b\|_\infty$ its supremum.

We assume that the initial conditions are independent and with a law that only depends on the heterogeneity parameters $(r_i)_{i=1,\cdots,N}$. This initial condition, that we call *chaotic*, corresponds to the assumption that there exists a collection of laws $(\mu_0(r))_{r\in D} \in \mathcal{M}_1^+(\mathbb{R})$ such that

$$\text{Law of } (x_0) = \bigotimes_{i=1}^{N} \mu_0(r_i). \tag{4}$$

The classical good assumptions on f and b ensures the well-posedness of the system:

Proposition 1. *For each $J \in \mathbb{R}^{N\times N}$, $\mathbf{r} \in D^N$, and $T > 0$, there exists a unique weak solution to the system* (1) *defined on $[0, T]$ with initial condition* (4)*. Moreover, this solution is square integrable.*

Let $T > 0$, and $Q_{\mathbf{r}}^N(J)$ be this unique weak solution restricted to the σ-algebra $\sigma(X_s^{i,N}, 1 \le i \le N, 0 \le s \le T)$. $Q_{\mathbf{r}}^N(J)$ is a probability measure on $\mathcal{C}^N$, where $\mathcal{C} := \mathcal{C}\big([0, T], \mathbb{R}\big)$, and depends on both realizations of J and $\mathbf{r}$. We are interested in proving a Large Deviation Principle (LDP) for the double-layer empirical measure,

$$\hat{\mu}_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{(X^{i,N},r_i)}. \tag{5}$$

In this perspective, let us introduce the system without interaction, i.e. when $J_{ij}=0$ for all (i,j). In this case, the law of a node with intrinsic dynamic with heterogeneity parameter $r\in D$ is given by the unique solution of this one-dimensional SDE:

$$\begin{cases} dX_t = f(r,t,X_t)dt + \lambda dW_t \\ (X_0) \overset{\mathcal{L}}{=} \mu_0(r). \end{cases} \tag{6}$$

We denote its restriction to the σ-algebra $\mathcal{G}_T=\sigma(X_s, s\le T)$ by P_r. As shown in [8, Appendix B], we can integrate this probability on the possible realizations of r to obtain a well-defined probability measure on $\mathcal{M}_1^+\big(C\times D\big)$ given by $dP(x,r) := dP_r(x)d\pi(r)$. Under $P^{\otimes N}$, particles are i.i.d. so that Sanov's theorem ensures the existence of a full LDP for the empirical measure, with good rate function given by the relative entropy $I(.|P)$[1]. The purpose of this article is to derive, from this result, another LDP for the interacting system.

A direct application of Girsanov theorem yields that $Q^N_{\mathbf{r}}(J)$ is absolutely continuous with respect to $P_{\mathbf{r}} := \bigotimes_{i=1}^N P_{r_i}$, with density:

$$\frac{\mathrm{d}Q^N_{\mathbf{r}}(J)}{\mathrm{d}P_{\mathbf{r}}}(\mathbf{x},\mathbf{r}) \;=\; \exp\left(\sum_{i=1}^{N}\int_0^T\Big(\frac{1}{\lambda}\sum_{j=1}^N J_{ij}b(x_t^i,x_t^j)\Big)dW_t(x^i,r_i) - \frac{1}{2}\int_0^T\Big(\frac{1}{\lambda}\sum_{j=1}^N J_{ij}b(x_t^i,x_t^j)\Big)^2 dt\right), \tag{7}$$

where $W_t(x,r) := \frac{x_t-x_0}{\lambda} - \int_0^t \frac{f(r,s,x_s)}{\lambda}ds,\ \forall (x,r)\in\mathcal{C}\times D$ in order to make $W(.,r)$ a Brownian motion under P_r. Moreover, as done in [8, Appendix B], we can define properly the averaged probability measure $Q^N := \mathcal{E}\big(Q^N_{\mathbf{r}}(J)\big)\in\mathcal{M}_1^+\big((\mathcal{C}\times D)^N\big)$. We now state our main results.

Theorem 2. *Under $Q^N=\mathcal{E}(Q^N_{\mathbf{r}}(J))$ the law of the N-particles system averaged over all possible configurations (realizations of $(J,\mathbf{r})$), $\hat{\mu}_N$ converges towards δ_Q.*

This convergence is the consequence of the following two theorems. The first provides a weak LDP for the empirical measure through an upper-bound for compact sets and the tightness of the sequence $Q^N\big(\hat{\mu}_N\in .\big)$. The second characterizes the unique minimum of the good rate function.

Theorem 3.
1. *There exists a good rate function $H:\mathcal{M}_1^+(\mathcal{C}\times D)$ such that for any compact subset K of $\mathcal{M}_1^+(\mathcal{C}\times D)$:*

$$\limsup_{N\to\infty}\frac{1}{N}\log Q^N(\hat{\mu}_N\in K)\le -\inf_K H.$$

2. *For any real number $\varepsilon>0$, there exists a compact subset K_ε such that for any integer N,*

$$Q^N(\hat{\mu}_N\notin K_\varepsilon)\le\varepsilon.$$

This theorem is proved in section 3.

Theorem 4. *The good rate function H is such that:*

[1] We recall that if Σ is a Polish space, the relative entropy of $\nu\in\mathcal{M}_1^+(\Sigma)$ with respect to μ is defined by:

$$I(\nu|\mu) := \begin{cases} \int \log\frac{d\nu}{d\mu}d\nu & \text{if } \nu\ll\mu, \\ \infty & \text{otherwise}. \end{cases}$$

1. *It achieves its minimal value at a unique probability measure $Q \in \mathcal{M}_1^+(\mathcal{C} \times D)$ satisfying:*

$$Q \simeq P, \qquad \frac{dQ}{dP}(x,r) = \mathcal{E}\left[\exp\left\{\frac{1}{\lambda}\int_0^T G_t^Q(x)dW_t(x,r) - \frac{1}{2\lambda^2}\int_0^T (G_t^Q(x))^2 dt\right\}\right]$$

where $(W_t(.,r))_{t\in[0,T]}$ is a P_r-brownian motion, and $G^Q(x)$ is, under $\mathcal{P}$, a Gaussian process with mean:

$$\mathcal{E}[G_t^Q(x)] = \int_{\mathcal{C}\times D} \bar{J} b(x_t, y_t) dQ(y, r')$$

and covariance:

$$\mathcal{E}[G_t^Q(x) G_s^Q(x)] = \int_{\left(\mathcal{C}\times D\right)^2} \sigma^2 b(x_t, y_t) b(x_s, y_s) dQ(y, r').$$

2. *This provides an implicit self-consistent equation on the limit distribution Q.*

This theorem will be demonstrated in section 4.

Based on this result, we can further conclude on the following:

Theorem 5. *For any connectivity matrix J, the system enjoys the propagation of chaos property. In other terms, Q^N is Q-chaotic, i.e. for any bounded continuous functions $(f_1, \cdots, f_m)$ and any indexes $(k_1, \cdots, k_m)$, we have:*

$$\lim_{N\to\infty} \int \prod_{j=1}^m f_j(x^{k_j}, r_{k_j}) dQ^N(x) = \prod_{j=1}^m \int f_j(x) dQ(x).$$

This is a direct consequence of theorem 2, thanks to a result due to Alain-Sol Sznitman, see [29, Lemma 3.1].

3 Large deviation principle

This section is devoted to proving the existence of a weak large deviations principle for the averaged empirical measure. We start by constructing the appropriate good rate function before proving the associated upper-bound and tightness results. Many point of the proof proceed as in precedent work [3, 18, 7, 8]. To avoid reproduce fastidious proofs, already known in the literature, we will, in the sequel, often rely on these precedent articles, and focus on the new difficulties arising from our setting.

3.1 Construction of the good rate function

For $\mu \in \mathcal{M}_1^+(\mathcal{C}\times D)$, we define the two following functions defined respectively on $[0,T]^2 \times \mathcal{C}$ and $[0,T]\times\mathcal{C}$:

$$\begin{cases} K_\mu(s,t,x) & := \dfrac{\sigma^2}{\lambda^2}\displaystyle\int_{\mathcal{C}} b(x_t, y_t) b(x_s, y_s) d\mu(y, r') \\ m_\mu(t,x) & := \dfrac{\bar{J}}{\lambda}\displaystyle\int_{\mathcal{C}} b(x_t, y_t) d\mu(y, r'). \end{cases}$$

Both functions are bounded: $|K_\mu(s,t,x)| \leq \frac{\sigma^2 \|b\|_\infty^2}{\lambda^2}$ and $|m_\mu(t,x)| \leq \frac{\bar{J}\|b\|_\infty}{\lambda}$. Moreover, as μ charges continuous functions, K_μ and m_μ are continuous maps by the dominated convergence theorem.

Since K_μ has a covariance structure, we can define a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \gamma)$ and a family of stochastic processes $\left(G^\mu(x)\right)_{x\in\mathcal{C},\mu\in\mathcal{M}_1^+(\mathcal{C}\times D)}$, independent for different x, and such that $G^\mu(x)$ is a γ-centered Gaussian process with covariance $K_\mu(.,.,x)$ (see [18, Remark 2.14]). We denote $\mathcal{E}_\gamma$ the expectation under γ.

Remark 1. Let $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$, and let $(e_i^\mu)_{i\in\mathbb{N}}$ be an orthonormal basis of $L^2_\mu(\mathcal{C} \times D)$. Let also for any $x \in \mathcal{C}, t \in [0,T]$, $\rho_{t,x} \in L^2_\mu$ such that $\rho_{t,x}(y,r) := b(x_t, y_t)$. As stated in [18], a possible explicit construction for the $G^\mu(x)$ is given by

$$G_t^\mu(x) := \sum_{i\in\mathbb{N}} J_i(x)\langle \rho_{t,x}, e_i^\mu\rangle_\mu = \sum_{i\in\mathbb{N}} J_i(x) \int b(x_t, y_t) e_i^\mu(y,r)\mu(y,r),$$

where the $\big(J_i(x)\big)_{i\in\mathbb{N},x\in\mathcal{C}}$ are independent centered Gaussian variables of variance σ^2.

We define

$$X^\mu(x,r) := \int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big)dW_t(x,r) - \frac{1}{2}\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big)^2 dt.$$

Similarly, we define for any fixed $N \in \mathbb{N}$

$$X^i(\mathbf{x},\mathbf{r}) := \int_0^T G_t^{i,N}(\mathbf{x})dW_t(x^i, r_i) - \frac{1}{2}\int_0^T G_t^{i,N}(\mathbf{x})^2 dt$$

with $G_t^{i,N}(\mathbf{x}) := \frac{1}{\lambda}\sum_{j=1}^N J_{ij} b(x_t^i, x_t^j)$.

We will use in the demonstration the following inequalities related to the relative entropy. For p and q two probability measures on a Polish space E, we recall the following identity (see e.g. [15, Lemma 3.2.13])

$$I(q|p) = \sup\left\{\int_E \Phi dq - \log\int_E \exp\Phi dp\ ;\ \Phi \in \mathcal{C}_b(E)\right\},$$

which implies in particular that for any bounded measurable function Φ on E,

$$\int_\mathcal{C} \Phi dq - \log\int_\mathcal{C} \exp\Phi dp \le I(q|p). \tag{8}$$

If Φ is a positive measurable function this inequality holds by monotone convergence, thus:

$$\int_\mathcal{C} \Phi dq \le I(q|p) + \log\int_\mathcal{C} \exp\Phi dp. \tag{9}$$

We know state a key result to our analysis:

Lemma 6.

$$\frac{dQ^N}{dP^{\otimes N}} = \exp\Big\{N\Gamma(\hat\mu_N)\Big\}.$$

where, for every $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$,

$$\Gamma(\mu) := \int_{\mathcal{C}\times D} \log\left(\mathcal{E}_\gamma\Big[\exp\big\{X^\mu(x,r)\big\}\Big]\right) d\mu(x,r). \tag{10}$$

Proof. Averaging the expression (7) on J and applying Fubini theorem, we find that $Q^N \ll P^{\otimes N}$ and, by independence of the J_{ij},

$$\frac{dQ^N}{dP^{\otimes N}}(\mathbf{x},\mathbf{r}) = \prod_{i=1}^N \mathcal{E}_J\Big[\exp\big(X^i(\mathbf{x},\mathbf{r})\big)\Big].$$

Here, $\mathbf{x}$ is the coordinate process taken under $P^{\otimes N}$, and displays no dependence with the J_{ij}. Moreover $\Big\{G_t^{i,N}(\mathbf{x}), 0 \le t \le T\Big\}$ is, under $\mathcal{P}_J$, a Gaussian process with covariance $K_{\hat\mu_N}(t,s,x^i)$, and mean $m_{\hat\mu_N}(t,x^i)$. We can thus write:

$$\frac{dQ^N}{dP^{\otimes N}} = \exp\left\{\sum_{i=1}^N \log\left(\mathcal{E}_\gamma\left[\exp\Big\{X^{\hat\mu_N}(x^i,r_i)\Big\}\right]\right)\right\} = \exp\Big\{N\Gamma(\hat\mu_N)\Big\}.$$

□

Γ has the following properties:

Proposition 7. *For every $\mu \in \mathcal{M}_1^+\big(\mathcal{C} \times D\big)$,*

1. *$\Gamma(\mu) \le I(\mu|P)$. In particular, $\Gamma(\mu)$ whenever $I(\mu|P)$ is.*
2. *If $\frac{2\sigma^2 \|b\|_\infty^2 T}{\lambda^2} < 1$, there exist real constants $\iota < 1$ and $\eta > 0$, such that $\Gamma(\mu) \le \iota I(\mu|P) + \eta$.*

Proof. **(i):**
Let F_μ denote the integrand in the formulation of Γ (10), and $F_{\mu,M}$:

$$F_{\mu,M}(x,r) := \log\Big\{\mathcal{E}_\gamma\Big[M \wedge \exp\Big\{X^\mu(x,r)\Big\}\Big]\Big\}.$$

The latter functional is positive bounded and measurable, thus inequality (9) holds. Taking M to infinity, the monotone convergence theorem ensures that, for every $a \ge 1$,

$$\begin{aligned} a\int_{\mathcal{C}} F_\mu(x) d\mu(x,r) &\le I(\mu|P) + \log\Big\{\int \exp aF_\mu(x) dP\Big\} = I(\mu|P) + \log\Big\{\int \mathcal{E}_\gamma\Big[\exp\Big\{X^\mu(x,r)\Big\}\Big]^a dP(x,r)\Big\} \\ &\overset{\text{Jensen}}{\le} I(\mu|P) + \log\Big\{\int \mathcal{E}_\gamma\Big[\exp\Big\{aX^\mu(x,r)\Big\}\Big] dP(x,r)\Big\} \overset{\text{Fubini}}{\le} I(\mu|P) + \log\Big\{\mathcal{E}_\gamma\Big[\int \exp\Big\{aX^\mu(x,r)\Big\} dP(x,r)\Big]\Big\}. \end{aligned} \tag{11}$$

$W(\cdot,r)$ being a P_r-Brownian motion, using the martingale property with $a=1$ completes the proof.

(ii):
Let $a > 1$. As $G^\mu(x)$ cannot be extracted from the integral on $dP_r(x)$, we rely on Hölder inequality with conjugate exponents $(a, \frac{a}{a-1})$ to make use of a martingale property.

$$\begin{aligned} &\int \mathcal{E}_\gamma\Big[\exp\Big\{a\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big) dW_t(x,r) - \frac{a}{2}\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big)^2 dt\Big\}\Big] dP(x,r) \le \\ &\quad \Big\{\int \mathcal{E}_\gamma\Big[\exp\Big\{a^2\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big) dW_t(x,r) - \frac{a^4}{2}\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big)^2 dt\Big\}\Big] dP(x,r)\Big\}^{\frac{1}{a}} \\ &\quad \times \Big\{\int \mathcal{E}_\gamma\Big[\exp\Big\{\frac{a^2(a+1)}{2}\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big)^2 dt\Big\}\Big] dP(x,r)\Big\}^{\frac{a-1}{a}}. \end{aligned}$$

The first term is equal to 1 by the martingale property. The short-time hypothesis $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$ ensures finiteness of the second term: indeed, by Jensen and Fubini inequalities, we have

$$\mathcal{E}_\gamma\Big[\exp\Big\{\frac{a^2(a+1)T}{2}\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big)^2 \frac{dt}{T}\Big\}\Big] \le \int_0^T \int \exp\Big\{\frac{a^2(a+1)T}{2}\big(G_t^\mu(x) + m_\mu(t,x)\big)^2\Big\}\Big]\frac{dt}{T}.$$

Moreover, since $\sqrt{a^2(a+1)T}\Big(G_t^\mu(x) + m_\mu(t,x)\Big) \sim \mathcal{N}\Big(\sqrt{a^2(a+1)T}m_\mu(t,x), a^2(a+1)TK_\mu(t,t,x)\Big)$ under γ, we are able, for $a-1$ small enough and under the short time hypothesis, to use the following identity, valid for $\zeta \sim \mathcal{N}(\alpha,\beta)$ with $\beta < 1$:

$$\mathbb{E}\Big[\exp\Big\{\frac{1}{2}\zeta^2\Big\}\Big] = \frac{1}{\sqrt{1-\beta}}\exp\Big\{\frac{\alpha^2}{2(1-\beta)}\Big\} = \exp\Big\{\frac{1}{2}\Big(\frac{\alpha^2}{1-\beta} - \log(1-\beta)\Big)\Big\}. \tag{12}$$

We thus conclude that as soon as $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$, there exists a constant c_T, uniform in $x \in \mathcal{C}$, such that:

$$\Big\{\int \mathcal{E}_\gamma\Big[\exp\Big\{\frac{a^2(a+1)}{2}\int_0^T \big(G_t^\mu(x) + m_\mu(t,x)\big)^2 dt\Big\}\Big] dP(x,r)\Big\}^{\frac{a-1}{a}} \le \exp\big\{(a-1)c_T\big\}.$$

Therefore, by (11):

$$\Gamma(\mu) \leq \frac{1}{a} I(\mu|P) + (a-1)c_T.$$

$\square$

Although Γ is neither bounded or continuous, preventing to rely on Varadhan's lemma, we will prove that the map

$$H(\mu) := \begin{cases} I(\mu|P) - \Gamma(\mu) & \text{if } I(\mu|P) < \infty, \\ \infty & \text{otherwise .} \end{cases}$$

is a good rate function associated with the desired weak large deviations principle of theorem (3). To do so, let us introduce a linearization of Γ depending on a parameter $\nu \in \mathcal{M}_1^+\big(\mathcal{C} \times D\big)$:

$$\Gamma_\nu(\mu) := \int_{\mathcal{C}} \log\left(\mathcal{E}_\gamma\left[\exp\left\{X^\nu(x,r)\right\}\right]\right) d\mu(x,r).$$

A key observation is that

$$\exp\left\{N\Gamma_\nu(\hat{\mu}_N)\right\} dP^{\otimes N}(\mathbf{x}) = \Big(\exp\left\{\Gamma_\nu(\delta_{(x,r)})\right\} dP(x,r)\Big)^{\otimes N} =: dQ_\nu(x,r)^{\otimes N}.$$

Sanov's theorem ensures that the empirical measure satisfies a full LDP under $Q_\nu^{\otimes N}$, with good rate function $I(.|Q_\nu)$. Shall Varadhan lemma apply, the good rate function would also be given by

$$H_\nu : \begin{cases} \mathcal{M}_1^+(\mathcal{C} \times D) & \to \mathbb{R}^+ \\ \mu & \to \begin{cases} I(\mu|P) - \Gamma_\nu(\mu) & \text{if } I(\mu|P) < \infty, \\ \infty & \text{otherwise .} \end{cases} \end{cases}$$

We can show as in [8, Theorem 11] this intuitive result:

Theorem 8. *Q_ν is a well defined probability measure on $\mathcal{M}_1^+(\mathcal{C} \times D)$, and $H_\nu(\mu) = I(\mu|Q_\nu)$. In particular H_ν is a good rate fonction.*

We introduce the Vaserstein distance on $\mathcal{M}_1^+(\mathcal{C} \times D)$, compatible with the weak topology:

$$d_T(\mu,\nu) := \inf_\xi \left\{ \int \sup_{0 \leq t \leq T} |x_t - y_t|^2 + |r - r'| d\xi\big((x,r),(y,r')\big) \right\}^{\frac{1}{2}}$$

the infimum being taken on the laws ξ with marginals μ and ν. This metric will control the error between H and its pleasant approximation H_ν:

Theorem 9. *1. $\exists C_T > 0$, such that for every $\mu, \nu \in \mathcal{M}_1^+\big(\mathcal{C} \times D\big)$,*

$$|\Gamma_\nu(\mu) - \Gamma(\mu)| \leq C_T\big(1 + I(\mu|P)\big) d_T(\mu,\nu).$$

2. If $\frac{2\sigma^2 \|b\|_\infty^2 T}{\lambda^2} < 1$, H is a good rate function.

Proof. The basic mechanism for the proof is similar to [2, Lemma 3.3-3.4]. However, the dependence in x of the Gaussian $G^\mu(x)$ introduces specific technicalities on which we focus our attention to handle point (i). Moreover, point (ii) previously shown without restriction on time in cases where $b(x,y)$ is independent of x is now shown under the short-time hypothesis of Proposition 7 (ii).

In detail, it is proven in [7, 8] that Γ_ν writes $\Gamma_\nu(\mu) = \Gamma_{1,\nu}(\mu) + \Gamma_{2,\nu}(\mu)$ with

$$\Gamma_{1,\nu}(\mu) := \int_{\mathcal{C}\times D} \left\{ \log\Big(\mathcal{E}_\gamma\Big[\exp\big(-\frac{1}{2}\int_0^T G_t^\nu(x)^2 dt\big)\Big]\Big) - \frac{1}{2}\int_0^T m_\nu(t,x)^2 dt \right\} d\mu(x,r),$$

$$\Gamma_{2,\nu}(\mu) := \frac{1}{2}\int\int\Big(\int G_t^\nu(x)(dW_t(x,r) - m_\nu(t,x)dt)\Big)^2 d\gamma_{\widetilde{K}^T_{\nu,x}} d\mu(x,r)$$
$$+\int\int m_\nu(t,x) dW_t(x,r) d\mu(x,r),$$

and

$$d\gamma_{\widetilde{K}^T_{\nu,x}} := \frac{\exp\Big\{-\frac{1}{2}\int_0^T \big(G_t^\nu(x)\big)^2 dt\Big\}}{\int \exp\Big\{-\frac{1}{2}\int_0^T \big(G_t^\nu(x)\big)^2 dt\Big\} d\gamma} d\gamma.$$

Moreover, in [3, Appendix A], the authors demonstrated that for any fixed $x \in \mathcal{C}$ and $\nu \in \mathcal{M}_1^+\big(\mathcal{C}\times D\big)$, $\gamma_{\widetilde{K}^T_{\nu,x}}$ is a probability measure on Ω under which $G^\nu(x)$ is a centered Gaussian process with covariance

$$\widetilde{K}^t_{\nu,x}(s,u) := \left(\int \frac{\exp\Big\{-\frac{1}{2}\int_0^t \big(G_u^\nu(x)\big)^2 du\Big\} G_u^\nu(x) G_s^\nu(x)}{\int \exp\Big\{-\frac{1}{2}\int_0^t \big(G_u^\nu(x)\big)^2 du\Big\} d\gamma} d\gamma\right).$$

The previous decomposition has the interest of splitting the difficulties: $|\Gamma_\nu(\mu) - \Gamma(\mu)| \leq |\Gamma_{1,\nu}(\mu) - \Gamma_1(\mu)| + |\Gamma_{2,\nu}(\mu) - \Gamma_2(\mu)|$. While the first term is easily controlled by $C_T d_T(\mu,\nu)$ (see [8, Lemma.12]), the semi-martingale term requires new tools. We will thereby restrict our proof to $|\Gamma_{2,\nu}(\mu) - \Gamma_2(\mu)| \leq C_T(1 + I(\mu|P))d_T(\mu,\nu)$.

Note that if $\mu \not\ll P$ the inequality is satisfied as $I(\mu|P) = \infty$. We can then suppose $\mu \ll P$. This implies that μ has a measurable density ρ_μ with respect to $\mathcal{B}\big(\mathcal{C}\times D\big)$:

$$d\mu(x,r) = \rho_\mu(x,r) dP(x,r) = \rho_\mu(x,r) dP_r(x) d\pi(r).$$

Hence, for $r \in D$ such that $c_\mu(r) := \int_D \rho_\mu(x,r) dP_r(x) \neq 0$, we can properly define $\mu_r \in \mathcal{M}_1^+(\mathcal{C})$ by $d\mu_r(x) := \frac{\rho_\mu(x,r)}{c_\mu(r)} dP_r(x)$. Of course $\mu_r \ll P_r$, and

$$d\mu(x,r) = d\mu_r(x) c_\mu(r) d\pi(r). \tag{13}$$

Remark that c_μ is a measurable function in space, and that the set $\{r \in D,\ c_\mu(r) = 0\}$ will not impact the value of the integral of interest.

Let ξ be a probability measure on $\big(\mathcal{C}\times D\big)^2$ with marginals μ and ν, and γ_ξ be the law of a bi-dimensional centered Gaussian process (G, G') with covariance K_ξ given by

$$K_\xi(s,t,x) := \frac{\sigma^2}{\lambda^2}\int \begin{pmatrix} b(x_s,y_s)b(x_t,y_t) & b(x_s,y_s)b(x_t,z_t) \\ b(x_s,z_s)b(x_t,y_t) & b(x_s,z_s)b(x_t,z_t) \end{pmatrix} d\xi\big((x,r'),(z,\tilde{r}')\big). \tag{14}$$

In the expression of $\Gamma_{2,\nu}(\mu)$ and $\Gamma_2(\mu)$ we can then replace the triplet (G^μ, G^ν, γ) by (G, G', γ_ξ). Let

$$\Lambda_T\big(G(x)\big) = \frac{\exp\Big(-\frac{1}{2}\int_0^T G_t(x)^2 dt\Big)}{\int \exp\Big(-\frac{1}{2}\int_0^T G_t(x)^2 dt\Big) d\gamma_\xi}.$$

Then,

$$|\Gamma_{2,\nu}(\mu)-\Gamma_2(\mu)| \leq \frac{1}{2}\Big|\int\int\Big(\int G'_t(x)(dW_t(x,r)-m_\nu(t,x)dt)\Big)^2\Big(\Lambda_T\big(G'(x)\big)-\Lambda_T\big(G(x)\big)\Big)d\gamma_\xi d\mu(x,r)\Big|$$

$$+\frac{1}{2}\Big|\int\!\!\int\Big\{\Big(\int G_t(x)(dW_t(x,r)-m_\mu(t,x)dt)\Big)^2-\Big(\int G'_t(x)(dW_t(x,r)-m_\nu(t,x)dt)\Big)^2\Big\}\Lambda_T\big(G(x)\big)d\gamma_\xi d\mu(x,r)\Big|$$

$$+\Big|\int\int(m_\nu-m_\mu)(t,x)dW_t(x,r)d\mu(x,r)\Big|$$

$$\overset{C.S.}{\leq}\overbrace{\frac{1}{2}\int\int\Big|\Lambda_T(G(x))-\Lambda_T(G'(x))\Big|\Big(\int_0^T G_t(x)(dW_t(x,r)-m_\nu(t,x)dt)\Big)^2 d\gamma_\xi d\mu(x,r)}^{B_1}$$

$$+\underbrace{\frac{1}{2}\prod_{\varepsilon=\pm1}\Big(\int\int\Lambda_T(G(x))\Big(\int_0^T\big(G_t(x)+\varepsilon G'_t(x)\big)(dW_t(x,r)-m_\nu(t,x)dt)\Big)^2 d\gamma_\xi d\mu\Big)^{\frac{1}{2}}}_{B_2}$$

$$+\underbrace{\frac{1}{2}\prod_{\varepsilon=\pm1}\Big(\int\int\Lambda_T(G(x))\Big\{\int_0^T G_t(x)\Big((1+\varepsilon)dW_t(x,r)-\big(m_\mu(t,x)+\varepsilon m_\nu(t,x)\big)dt\Big)\Big\}^2 d\gamma_\xi d\mu\Big)^{\frac{1}{2}}}_{B_3}$$

$$+\underbrace{\Big(\int\Big|\int_0^T(m_\nu-m_\mu)(t,x)dW_t(x,r)\Big|^2 d\mu\Big)^{\frac{1}{2}}}_{B_4}.$$

Remark that these four terms are of the form

$$\int\int F(G,G')(x)\Big(\int_0^T H_t(G,G',\mu,\nu)(x)\big(\alpha dW_t(x,r)-M_t(\mu,\nu)(x)dt\big)\Big)^2 d\gamma_\xi d\mu(x,r)$$

with α equals 0 or 1. Controlling them will be the object of the following technical lemma.

Lemma 10. *For $\mu,\nu\in\mathcal{M}_1^+(\mathcal{C}\times D)$, let*

$$A(\mu,\nu):=\int\int F(G,G')(x)\Big(\int_0^T H_t(G,G',\mu,\nu)(x)\big(\alpha dW_t(x,r)-M_t(\mu,\nu)(x)dt\big)\Big)^2 d\gamma_\xi d\mu(x,r)$$

where (G,G') are, under γ_ξ, centered Gaussian processes with covariance given by (14), H and M are continuous progressively measurable processes, F is a positive measurable function, and $\alpha\in\{0,1\}$. Then,

$$A(\mu,\nu)\leq C_T\Big(\alpha\big(I(\mu|P)+1\big)\sup_{\mathcal{C}\times[0,T]}\Big\{\mathcal{E}_\xi\Big[H_t^4F^2\Big]\Big\}^{\frac{1}{2}}+\alpha\sup_{\mathcal{C}\times[0,T]}\Big\{\mathcal{E}_\xi\Big[FH_t^2\Big]\Big\}+\sup_{\mathcal{C}\times[0,T]}\Big\{M_t^2\mathcal{E}_\xi\Big[FH_t^2\Big]\Big\}\Big).\quad(15)$$

Proof.

$$A(\mu,\nu)\leq 2\alpha\int\int F\Big(\underbrace{\int_0^T H_t dW_t}_{N_T}\Big)^2 d\gamma_\xi d\mu+2\int\int F\Big(\int_0^T H_tM_tdt\Big)^2 d\gamma_\xi d\mu(x,r).$$

Ito calculus gives

$$N_T^2=2\int_0^T H_tN_tdW_t+\int_0^T H_t^2dt,$$

so that.

$$A(\mu,\nu) \le 4\int \underbrace{\int_0^T \mathcal{E}_\xi\Big[\alpha H_t F N_t\Big] dW_t}_{\tilde{N}_T} d\mu + 2\int\int_0^T \mathcal{E}_\xi\Big[\alpha F H_t^2\Big] dt d\mu + 2T\int\int_0^T M_t^2 \mathcal{E}_\xi\Big[F H_t^2\Big] dt d\mu, \quad (16)$$

The two last terms of the righthand side of (16) are easy to handle taking the supremum on $\mathcal{C}\times[0,T]$. The first one vanishes if $\alpha=0$, but is tricky when $\alpha=1$. In this last case, inequality (9) brings

$$\int \tilde{N}_T d\mu \overset{C.S}{\le} 2\Big(\int \langle \tilde{N}\rangle_T d\mu\Big)^{\frac{1}{2}}\Big(I(\mu|P) + \log\Big\{\int \exp\big\{\frac{\tilde{N}_T^2}{4\langle \tilde{N}\rangle_T}\big\} dP\Big\}\Big)^{\frac{1}{2}}.$$

As $\tilde{N}$ is a P-local martingale, Dambis-Dubins-Schwarz theorem ensures that $\frac{\tilde{N}_T^2}{4\langle \tilde{N}\rangle_T}$ has the same law as $\frac{B^2_{\langle \tilde{N}\rangle_T}}{4\langle \tilde{N}\rangle_T}$, where B is some P-Brownian Motion, so that exists a universal constant C satisfying

$$\log\Big\{\int \exp\big\{\frac{\tilde{N}_T^2}{4\langle \tilde{N}\rangle_T}\big\} dP\Big\} \le C.$$

Hence

$$\begin{aligned}\int \tilde{N}_t d\mu &\le \tilde{C}\Big(\int\int_0^T \mathcal{E}_\xi\Big[H_t F N_t\Big]^2 dt d\mu\Big)^{\frac{1}{2}}\big(I(\mu|P)+1\big)^{\frac{1}{2}} \\ &\overset{\text{C.S.}}{\le} \tilde{C}\sup_{x\in\mathcal{C}}\Big\{\mathcal{E}_\xi\Big[\langle N\rangle_t H_t^2 F^2\Big]\Big\}^{\frac{1}{2}}\Big(\int_0^T \mathcal{E}_\xi\Big[\int \frac{N_t^2}{4\langle N\rangle_t} d\mu\Big] dt\Big)^{\frac{1}{2}}\big(I(\mu|P)+1\big)^{\frac{1}{2}} \\ &\le C_T \sup_{\mathcal{C}\times[0,T]}\Big\{\mathcal{E}_\xi\Big[H_s^2 H_t^2 F^2\Big]\Big\}^{\frac{1}{2}}\big(I(\mu|P)+1\big).\end{aligned}$$

Cauchy-Schwarz inequality then yields the result.

□

To bound our four terms, we only have to check the majoration of the quantities $\alpha\mathcal{E}_\xi\Big[H_t^4 F^2\Big]$, $\alpha\mathcal{E}_\xi\Big[F H_t^2\Big]$, and $M_t^2\mathcal{E}_\xi\Big[F H_t^2\Big]$ by the Vaserstein distance times a uniform constant on $\mathcal{C}\times[0,T]$. We will prove this for the first term, as an example, and leave the three other terms as an exercice to the involved readers. Remark that,

$$\Lambda_T(G(x)) = \exp\Big\{-\log\Big(\int \exp\big(-\frac{1}{2}\int_0^T G_t(x)^2 dt\big) d\gamma_\xi\Big) - \frac{1}{2}\int_0^T G_t(x)^2 dt\Big\},$$

so that we can apply Jensen inequality to obtain

$$\Lambda_T(G(x)) \le \exp\Big\{\frac{\sigma^2\|b\|_\infty^2 T}{2\lambda^2}\Big\}.$$

This yields

$$\begin{aligned}\big|\Lambda_T(G(x)) - \Lambda_T(G'(x))\big| \le \exp\Big\{\frac{\sigma^2\|b\|_\infty^2 T}{2\lambda^2}\Big\}\Big(&\underbrace{\frac{1}{2}\int_0^T \big|G_t(x)^2 - G'_t(x)^2\big| dt}_{F_1} + \\ &\underbrace{\Big|\log\Big(\int \exp\big(-\frac{1}{2}\int_0^T G_t(x)^2 dt\big) d\gamma_\xi\Big) - \log\Big(\int \exp\big(-\frac{1}{2}\int_0^T G'_t(x)^2 dt\big) d\gamma_\xi\Big)\Big|}_{F_2}\Big).\end{aligned} \quad (17)$$

Moreover,

$$F_2(x) = \Big|\log\Big(\frac{\mathcal{E}_\xi\Big[\exp\big\{-\frac{1}{2}\int_0^T G_t(x)^2dt\big\}\Big]}{\mathcal{E}_\xi\Big[\exp\big\{-\frac{1}{2}\int_0^T G'_t(x)^2dt\big\}\Big]}\Big)\Big\}\Big| \leq \Big|\log\Big(1+\frac{\mathcal{E}_\xi\Big[\frac{1}{2}\int_0^T\big|G_t(x)^2-G'_t(x)^2\big|dt\big\}\Big]}{\mathcal{E}_\xi\Big[\exp\big\{-\frac{1}{2}\int_0^T G'_t(x)^2dt\big\}\Big]}\Big)\Big\}\Big|$$
$$\leq \frac{1}{2}\exp\Big\{\frac{\sigma^2\|b\|_\infty^2 T}{2\lambda^2}\Big\}\mathcal{E}_\xi\Big[\int_0^T\big|G_t(x)^2-G'_t(x)^2\big|dt\Big] \overset{C.S.}{\leq} C_T\mathcal{E}_\xi\Big[\int_0^T\big(G_t(x)-G'_t(x)\big)^2dt\Big]^{\frac{1}{2}}$$
$$\leq C_T\Big\{\int\int_0^T\big(b(x_s,y_s)-b(x_s,z_s)\big)^2dtd\xi\big((y,r'),(z,\tilde{r}')\big)\Big\}^{\frac{1}{2}} \leq C_T d_T(\mu,\nu) \tag{18}$$

so that one has for $p\in\{1,2\}$,

$$\mathcal{E}_\xi\Big[G_t^{2p}(x)F_2(x)\Big] \leq \mathcal{E}_\xi\Big[G_t^{2p}(x)\Big]C_T d_T(\mu,\nu) \leq C_T\frac{\sigma^{2p}\|b\|_\infty^{2p}}{\lambda^{2p}}d_T(\mu,\nu)$$

as for

$$\mathcal{E}_\xi\Big[G_t^2F_2\Big]m_\nu(t,x) \leq C_T\frac{\sigma^2\|b\|_\infty^3\bar{J}}{\lambda^3}d_T(\mu,\nu),$$

where both upper-bounds is uniform in (x,t). Isserlis' Theorem allows showing similar upper-bounds for the terms involving F_1.

□

3.2 Upper-bound and Tightness

We prove here a weak LDP relying on an upper-bound inequality for compact subsets, and tightness of the family $\Big(Q^N\big(\hat{\mu}_N\in .\big)\Big)_N$. To prove the first point, we take advantage of the full LDP followed by $\hat{\mu}_N$ under $\big(Q_\nu\big)^{\otimes N}$, and control an error. The second point will rely on the exponential tightness of $P^{\otimes N}$. These proofs are very close to the one developed by Guionnet [18].

Theorem 11. *Under the condition* $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2}<1$, *we have:*

1. *For any compact subset* K *of* $\mathcal{M}_1^+(\mathcal{C}\times D)$,

$$\limsup_{N\to\infty}\frac{1}{N}\log Q^N(\hat{\mu}_N\in K) \leq -\inf_K H.$$

2. *For any real number* $\varepsilon>0$, *there exists a compact set* K_ε *of* $\mathcal{M}_1^+(\mathcal{C}\times D)$ *such that, for any integer* N,

$$Q^N(\hat{\mu}_N\notin K_\varepsilon)\leq\varepsilon.$$

Proof. **(1):** Let $\delta<0$. We can find an integer M and a family $(\nu_i)_{1\leq i\leq M}$ of $\mathcal{M}_1^+(\mathcal{C}\times D)$ such that

$$K\subset\bigcup_{i=1}^M B(\nu_i,\delta),$$

where $B(\nu_i,\delta)=\big\{\mu|d_T(\mu,\nu_i)<\delta\big\}$. A very classical result (see e.g. [12, lemma 1.2.15]), ensures that

$$\limsup\frac{1}{N}\log Q^N(\hat{\mu}_N\in K) \leq \max_{1\leq i\leq p}\limsup\frac{1}{N}\log Q^N(\hat{\mu}_N\in K\cap B(\nu_i,\delta)).$$

Lemma 10 yields:

$$
\begin{aligned}
Q^N(\hat{\mu}_N \in K \cap B(\nu,\delta)) &= \int_{\hat{\mu}_N \in K \bigcap B(\nu,\delta)} \exp\Big\{N\Gamma(\hat{\mu}_N)\Big\} dP^{\otimes N} \\
&= \int_{\hat{\mu}_N \in K \bigcap B(\nu,\delta)} \exp\Big\{N\big(\Gamma(\hat{\mu}_N) - \Gamma_\nu(\hat{\mu}_N)\big)\Big\} \exp\Big\{N\Gamma_\nu(\hat{\mu}_N)\Big\} dP^{\otimes N}.
\end{aligned}
$$

Observe that, for conjugate exponents (p,q),

$$
\begin{aligned}
Q^N(\hat{\mu}_N \in K \cap B(\nu,\delta)) &= \int_{\hat{\mu}_N \in K \bigcap B(\nu,\delta)} \exp\Big\{N\big(\Gamma(\hat{\mu}_N) - \Gamma_\nu(\hat{\mu}_N)\big)\Big\} dQ_\nu^{\otimes N} \\
&\leq Q_\nu^{\otimes N}\big(\hat{\mu}_N \in K \cap B(\nu,\delta)\big)^{\frac{1}{p}} \left(\int_{\hat{\mu}_N \in K \bigcap B(\nu,\delta)} \exp\Big\{qN\big(\Gamma(\hat{\mu}_N) - \Gamma_\nu(\hat{\mu}_N)\big)\Big\} dQ_\nu^{\otimes N}\right)^{\frac{1}{q}},.
\end{aligned}
\tag{19}
$$

Then, by definitions of Γ and Γ_ν:

$$
\begin{aligned}
&\int_{\hat{\mu}_N \in K \bigcap B(\nu,\delta)} \exp\Big\{qN\big(\Gamma(\hat{\mu}_N) - \Gamma_\nu(\hat{\mu}_N)\big)\Big\} dQ_\nu^{\otimes N} = \int_{\hat{\mu}_N \in K \cap B(\nu,\delta)} \left(\prod_{i=1}^N \frac{\mathcal{E}_\gamma\Big[\exp X^{\hat{\mu}_N}(x^i,r_i)\Big]}{\mathcal{E}_\gamma\Big[\exp X^\nu(x^i,r_i)\Big]}\right)^q dQ_\nu^{\otimes N} \\
&= \int_{\hat{\mu}_N \in K \cap B(\nu,\delta)} \mathcal{E}_\gamma\left[\prod_{i=1}^N \left(\exp\Big(X^{\hat{\mu}_N}(x^i,r_i) - X^\nu(x^i,r_i)\Big)\right) \frac{\prod_{i=1}^N \exp X^\nu(x^i,r_i)}{\mathcal{E}_\gamma\Big[\prod_{i=1}^N \exp X^\nu(x^i,r_i)\Big]}\right]^q dQ_\nu^{\otimes N} \\
&\overset{\text{Jensen}}{\leq} \Bigg\{\underbrace{\int_{\hat{\mu}_N \in K \cap B(\nu,\delta)} \mathcal{E}_\gamma\left[\prod_{i=1}^N \left(\exp q\Big(X^{\hat{\mu}_N}(x^i,r_i) - X^\nu(x^i,r_i)\Big) \exp X^\nu(x^i,r_i)\right)\right] dP^{\otimes N}}_{B_N}\Bigg\}^{\frac{1}{q}},
\end{aligned}
$$

so that

$$
Q^N(\hat{\mu}_N \in K \cap B(\nu,\delta)) \leq Q_\nu^{\otimes N}\big(\hat{\mu}_N \in K \cap B(\nu,\delta)\big)^{\frac{1}{p}} B_N^{\frac{1}{q}}
$$

The first term of the righthand side can be controlled by large deviations estimates. The second term boundness ensue from the following lemma 12, proven in Appendix 5.

Lemma 12. *For any real number $q > 1$, and if $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$, there exists a strictly positive real number δ_q such that, for any $\delta < \delta_q$, there exists a function $C_q(.)$ in $\mathbb{R}$ such that $\lim_{\delta\to 0} C_q(\delta) = 0$ and:*

$$
B_N \leq \exp\{C_q(\delta)N\}.
$$

Concluding the proof of the first point can now be done exactly as in [3, Lemma 4.7].

Proof of (2):

The proof of this theorem consists in using the relative entropy inequality (9) and the exponential tightness of the sequence $(P^{\otimes N})_N$. The reader shall refer to [7, Theorem 2], to obtain the inequality:

$$
\begin{aligned}
I(Q^N|P^{\otimes N}) &= \frac{N}{2} \int_{(\mathcal{C}\times D)^N} \int_0^T \left(\int_0^t \widetilde{K}^t_{\hat{\mu}_N,x^1}(t,s)\Big(dW_s(x^1,r_1) - m_{\hat{\mu}_N}(s,x^1)ds\Big) + m_{\hat{\mu}_N}(t,x^1)\right)^2 dt dQ^N(\mathbf{x},\mathbf{r}) \\
&\leq N \times \Bigg\{\int_0^T \underbrace{\int_{(\mathcal{C}\times D)^N} \left(\int_0^t \widetilde{K}^t_{\hat{\mu}_N,x^1}(t,s)\Big(dW_s(x^1,r_1) - m_{\hat{\mu}_N}(s,x^1)ds\Big)\right)^2 dQ^N(\mathbf{x},\mathbf{r})}_{\phi(t,\mathbf{x},\mathbf{r})} dt + \frac{\bar{J}^2 T}{\lambda^2}\Bigg\}.
\end{aligned}
$$

We then bound $\phi(t, \mathbf{x}, \mathbf{r})$ uniformly in space to conclude:

$$\sup_{t \leq T} \phi(t, \mathbf{x}, \mathbf{r}) \leq 2\frac{\sigma^4 \|b\|_\infty^4 T}{\lambda^4} \exp\left\{2\frac{\sigma^4 \|b\|_\infty^4 T}{\lambda^4}\right\}.$$

□

4 Existence and characterization of the limit

4.1 Uniqueness of the minimum

In this section, we rely on an original contraction argument to show that the good rate function H admits a unique minimum. By a variational approach, we have shown in [7] that any minimum Q of the good rate function H satisfies:

$$Q \simeq P, \qquad \frac{\mathrm{d}Q}{\mathrm{d}P}(x,r) = \mathcal{E}_\gamma\Big[\exp\left\{X^Q(x,r)\right\}\Big]. \tag{20}$$

The map $(x,r) \to \mathcal{E}_\gamma\Big[\exp\Big\{X^\mu(x,r)\Big\}\Big]$ is non-negative and measurable for every $\mu \in \mathcal{M}_1^+(\mathcal{C} \times D)$. Moreover, it is continuous as function of $m_\mu(.,x)$, $K_\mu(.,.,x)$ and $W_.(x,r)$. Hence, we can properly define

$$L := \begin{cases} \mathcal{M}_1^+\big(\mathcal{C} \times D\big) \to \mathcal{M}_1^+\big(\mathcal{C} \times D\big) \\ \mu \to dL(\mu)(x,r) := \mathcal{E}_\gamma\Big[\exp\Big\{X^\mu(x,r)\Big\}\Big] dP(x,r). \end{cases}$$

In fact, as $\exp\Big\{X^\mu(x,r)\Big\}$ is γ-almost surely finite, one can use Novikov criterion to show that $L(\mu)$ defines a probability measure on $\mathcal{C} \times D$. Equation (20) can be reformulated as follow: any minimum of H must satisfy

$$Q \simeq P, \qquad Q = L(Q).$$

Remark 2. *$L(\mu)$ is exactly Q_μ introduced for Theorem 8.*

Theorem 13. *The map L admits a unique fixed point.*

Proof. As in [3, Lemma 5.15], we can show that

$$\frac{\mathrm{d}L(\mu)}{\mathrm{d}P}(x,r) = \exp\left\{\int_0^T H_\mu(t,x,r) dW_t(x,r) - \frac{1}{2}\int_0^T H_\mu^2(t,x,r)dt\right\},$$

where

$$H_\mu(t,x,r) = \int_0^t \widetilde{K}_{\mu,x}^t(t,s)\big(dW_s(x,r) - m_\mu(s,x)ds\big) + m_\mu(t,x).$$

Let $\mu \in \mathcal{M}_1^+\big(\mathcal{C} \times D\big), r \in D$, and remark that $x \to \frac{\mathrm{d}L(\mu)}{\mathrm{d}P}(x,r)$ is a P_r-martingale. Hence one can properly define $L(\mu)_r \in \mathcal{M}_1^+(\mathcal{C})$ and use Girsanov theorem to remark that it is the unique weak solution of

$$\begin{cases} dx_t = f(r,t,x_t)dt + H_\mu^{\tilde{W}}(t,x_t,r)dt + \lambda(r)d\tilde{W}_t \\ x_0 = \bar{x}_0. \end{cases}$$

where $\tilde{W}$ is a $\mathbb{P}$-Brownian motion, $H_\mu^{\tilde{W}}(t,x,r) := \int_0^t \widetilde{K}_{\mu,x}^t(t,s)\big(d\tilde{W}_s - m_\mu(s,x)ds\big) + m_\mu(t,x)$, and $\bar{x}_0 \in \mathbb{R}$ is the realization of $\mu_0(r)$. We denote its unique strong solution by $\big(x_r^\mu(t)\big)_{t\in[0,T]}$. Let also $\nu \in \mathcal{M}_1^+\big(\mathcal{C} \times D\big)$, and define similarly x_r^ν with same initial condition and Brownian motion. In what follow, we will unambiguously drop the space index for our strong solutions, and the Brownian exponent for $H^{\tilde{W}}$. We have

$$\big(x_t^\mu - x_t^\nu\big) = \int_0^t \big(f(r,s,x_s^\mu) + m_\mu(s,x^\mu) - f(r,s,x_s^\nu) - m_\nu(s,x^\nu)\big)ds$$
$$+ \int_0^t \int_0^s \Big(\widetilde{K}^s_{\mu,x^\mu}(s,u)m_\mu(u,x^\mu) - \widetilde{K}^s_{\nu,x^\nu}(s,u)m_\nu(u,x^\nu)\Big)duds + \int_0^t \int_0^s \big(\widetilde{K}^s_{\mu,x^\mu}(s,u) - \widetilde{K}^s_{\nu,x^\nu}(s,u)\big)d\tilde{W}_u ds. \tag{21}$$

First, remark that

$$\begin{aligned} m_\mu(t,x^\mu) - m_\nu(t,x^\nu) &= \int_{\mathcal{C}\times D} b(x_t^\mu,y_t) - b(x_t^\nu,y_t)d\mu(y) + \int_{\mathcal{C}\times D} b(x_t^\nu,y_t) - b(x_t^\nu,z_t)d\xi\big((y,r'),(z,\tilde{r})')\big) \\ &\leq K_b\big|x_t^\mu - x_t^\nu\big| + K_b \int_{\mathcal{C}\times D} \sup_{s\leq t}\big|y_s - z_s\big|d\xi \end{aligned}$$

for any $\xi \in \mathcal{M}_1^+\big((\mathcal{C}\times D)^2\big)$ with marginals μ and ν. Let, for any $\chi \in \mathcal{M}_1^+\big(\mathcal{C}\times D\big), t \leq T$

$$\Lambda_t\big(G^\chi(x)\big) = \frac{\exp\Big(-\frac{1}{2}\int_0^t G_s^\chi(x)^2 ds\Big)}{\int \exp\Big(-\frac{1}{2}\int_0^t G_s^\chi(x)^2 ds\Big)d\gamma} d\gamma.$$

and remark that Jensen inequality gives $0 \leq \Lambda \leq \exp\big\{\frac{\sigma^2\|b\|_\infty^2 T}{2\lambda^2}\big\}$. Then

$$\begin{aligned} &\widetilde{K}^t_{\mu,x^\mu}(t,s) - \widetilde{K}^t_{\nu,x^\mu}(t,s) = \mathcal{E}_\gamma\Big(G_t^\mu(x^\mu)G_s^\mu(x^\mu)\Lambda_t\big(G^\mu(x^\mu)\big) - G_t^\nu(x^\mu)G_s^\nu(x^\mu)\Lambda_t\big(G^\nu(x^\mu)\big)\Big) \\ &= \mathcal{E}_\gamma\Big[\big(G_t^\mu(x^\mu) - G_t^\nu(x^\mu)\big)G_s^\mu(x^\mu)\Lambda_t\big(G^\mu(x^\mu)\big)\Big] + \mathcal{E}_\gamma\Big[G_t^\nu(x^\mu)\big(G_s^\mu(x^\mu) - G_s^\nu(x^\mu)\big)\Lambda_t\big(G^\mu(x^\mu)\big)\Big] \\ &+ \mathcal{E}_\gamma\Big[G_t^\nu(x^\mu)G_s^\nu(x^\mu)\Big(\Lambda_t\big(G^\mu(x^\mu)\big) - \Lambda_t\big(G^\nu(x^\mu)\big)\Big)\Big]. \end{aligned}$$

On one hand, Cauchy-Schwartz Theorem yields

$$\begin{aligned} \mathcal{E}_\gamma\Big[\big(G_t^\mu(x^\mu) - G_t^\nu(x^\mu)\big)G_s^\mu(x^\mu)\Lambda_t\big(G^\mu(x^\mu)\big)\Big] &\leq \mathcal{E}_\gamma\Big[\big(G_t^\mu(x^\mu) - G_t^\nu(x^\mu)\big)^2\Big]^{\frac{1}{2}} \exp\big\{\frac{\sigma^2\|b\|_\infty^2 T}{4\lambda^2}\big\}\sqrt{\widetilde{K}^t_{\mu,x^\mu}(s,s)} \\ &\leq C_T\Big\{\int_{\mathcal{C}\times D}\Big(b\big(x_t^\mu,y_t\big) - b\big(x_t^\mu,z_t\big)\Big)^2 d\xi\Big\}^{\frac{1}{2}} \leq C_T\Big\{\int_{\mathcal{C}\times D}\sup_{s\leq t}\big|y_s - z_s\big|^2 d\xi\Big\}^{\frac{1}{2}}. \end{aligned}$$

On the other hand,

$$\Lambda_t(G^\mu(x)) = \exp\Big\{-\log\Big(\int \exp\big\{-\frac{1}{2}\int_0^t G_s^\mu(x)^2 ds\big\}d\gamma\Big) - \frac{1}{2}\int_0^t G_s^\mu(x)^2 ds\Big\}$$

so that using (17) and (18), we obtain

$$\Big|\Lambda_t\big(G^\mu(x^\mu)\big) - \Lambda_t\big(G^\nu(x^\mu)\big)\Big| \leq C_T\Big\{\int_0^t \big|G_s^\mu(x^\mu)^2 - G_s^\nu(x^\mu)^2\big|ds + \mathcal{E}_\gamma\Big[\int_0^t \big|G_s^\mu(x^\mu)^2 - G_s^\nu(x^\mu)^2\big|ds\big\}\Big]\Big\}.$$

As a consequence, we can use Cauchy-Schwartz inequality and Isserlis' theorem to obtain

$$\mathcal{E}_\gamma\Big[G_t^\nu(x^\mu)G_s^\nu(x^\mu)\Big(\Lambda_t\big(G^\mu(x^\mu)\big) - \Lambda_t\big(G^\nu(x^\mu)\big)\Big)\Big] \leq C_T\Big\{\int_{\mathcal{C}\times D}\sup_{s\leq t}\big|y_s - z_s\big|^2 d\xi\Big\}^{\frac{1}{2}},$$

therefore yielding

$$\widetilde{K}^t_{\mu,x^\mu}(t,s) - \widetilde{K}^t_{\nu,x^\mu}(t,s) \leq C_T \Big\{ \int_{\mathcal{C}\times D} \sup_{u\leq t} \big|y_u - z_u\big|^2 d\xi \Big\}^{\frac{1}{2}}.$$

The exact same proof applies to obtain

$$\widetilde{K}^t_{\mu,x^\mu}(t,s) - \widetilde{K}^t_{\mu,x^\nu}(t,s) \leq C_T \sup_{u\leq t} \big|x^\mu_u - x^\nu_u\big|.$$

Hence,

$$\Big|\widetilde{K}^t_{\mu,x^\mu}(t,s) m_\mu(s,x^\mu) - \widetilde{K}^t_{\nu,x^\nu}(t,s) m_\nu(s,x^\nu)\Big|^2 \leq C_T \Big\{ \int_{\mathcal{C}\times D} \sup_{u\leq t} \big|y_u - z_u\big|^2 d\xi + \sup_{u\leq t} \big|x^\mu_u - x^\nu_u\big|^2 \Big\}.$$

Injecting these result in (21) writes

$$\sup_{s\leq t} \big|x^\mu_s - x^\nu_s\big|^2 \leq C_T \Big\{ \int_0^t \sup_{u\leq s} \big|x^\mu_u - x^\nu_u\big|^2 ds + \int_0^t \sup_{v\leq s} \Big(\int_0^v \big(\widetilde{K}^v_{\mu,x^\mu}(v,u) - \widetilde{K}^v_{\nu,x^\nu}(v,u)\big) d\tilde{W}_u \Big)^2 ds$$
$$+ \int_0^t \int_{\mathcal{C}\times D} \sup_{u\leq s} \big|y_u - z_u\big|^2 d\xi ds \Big\}.$$

Burkholder-Davis-Gundy inequality then ensures

$$\underbrace{\mathbb{E}\Big[\int_D \sup_{s\leq t} \big|x^\mu_r - x^\nu_r\big|^2(s) d\pi(r) \Big]}_{g(t)} \leq C_T \Big\{ \int_0^t g(s) ds + \int_0^t \int_{\mathcal{C}\times D} \sup_{u\leq s} \big|y_u - z_u\big|^2 d\xi ds \Big\}.$$

One can now use Gronwall lemma's and take the infimum on ξ to obtain

$$d_t\big(L(\mu), L(\nu)\big) \leq C_T \int_0^t d_s\big(\mu,\nu\big) ds.$$

The proof now relies on classical arguments. □

4.1.1 Convergence of the process

We are now in a position to prove theorem 2.

Proof of Theorem 2. Indeed, for δ a strictly positive real number and $B(Q,\delta)$ the open ball of radius δ centered in Q for the Vaserstein distance. We prove that $Q^N(\hat{\mu}_N \notin B(Q,\delta))$ tends to zero as N goes to infinity. Indeed, for K_ε a compact defined in theorem 11, we have for any $\varepsilon > 0$:

$$Q^N(\hat{\mu}_N \notin B(Q,\delta)) \leq \varepsilon + Q^N(\hat{\mu}_N \in B(Q,\delta)^c \cap K_\varepsilon).$$

The set $B(Q,\delta)^c \cap K_\varepsilon$ is a compact, and theorem 11 now ensures that

$$\limsup_{N\to\infty} \frac{1}{N} \log Q^N(\hat{\mu}_N \in B(Q,\delta)^c \cap K_\varepsilon) \leq - \inf_{B(Q,\delta)^c \cap K_\varepsilon} H$$

and eventually, theorem 4 ensures that the righthand side of the inequality is strictly negative, which implies that

$$\lim_{N\to\infty} Q^N(\hat{\mu}_N \notin B(Q,\delta)) \leq \varepsilon,$$

that is:

$$\lim_{N\to\infty} Q^N(\hat{\mu}_N \notin B(Q,\delta)) = 0.$$

□

4.1.2 Characterization of the limit

By Girsanov theorem, (20) implies that

$$d\bar{X}_t(r) = \Big(f\big(r,t,\bar{X}_t(r)\big) + G_t^{\bar{X}}\big(\bar{X}(r)\big)\Big)\,dt + \lambda dW_t(r) \tag{22}$$

where the processes $\big(W(r)\big)_r$, and $\big(G_t^{\bar{X}}(x)\big)_x$, are independent Brownian motions, Gaussian processes respectively. The latter have mean

$$m(t,x) = \bar{J}\mathbb{E}_{\bar{Z}}\Big[b\big(x_t,\bar{Z}_t\big)\Big]$$

and covariance

$$C(t,s,x) = \sigma^2\mathbb{E}_{\bar{Z}}\Big[b\big(x_s,\bar{Z}_s\big)b\big(x_t,\bar{Z}_t\big)\Big],$$

where $\bar{Z}$ denotes an independent copy of $\bar{X}$, and $\mathbb{E}_{\bar{Z}}$ is the expectation over $\bar{Z}$.

5 Appendix

Proof of Lemma 12:

Proof. Recall that

$$B_N = \int_{\hat{\mu}_N\in K\cap B(\nu,\delta)} \mathcal{E}\left[\prod_{i=1}^N \Big(\exp q\Big(X^{\hat{\mu}_N}(x^i,r_i) - X^\nu(x^i,r_i)\Big)\exp X^\nu(x^i,r_i)\Big)\right]dP^{\otimes N}.$$

Using Holder inequality with conjugate exponents (ρ,η), one finds:

$$B_N \le \Bigg\{\overbrace{\int\prod_{i=1}^N\mathcal{E}\Big[\exp\rho X^\nu(x^i,r_i)\Big]dP^{\otimes N}}^{B_1^N}\Bigg\}^{\frac{1}{\rho}}\Bigg\{\underbrace{\int_{\hat{\mu}_N\in B(\nu,\delta)}\mathcal{E}\Big[\prod_{i=1}^N\exp q\eta\Big(X^{\hat{\mu}_N}(x^i,r_i) - X^\nu(x^i,r_i)\Big)\Big]dP^{\otimes N}}_{B_2^N}\Bigg\}^{\frac{1}{\eta}} \tag{23}$$

On one hand, no new difficulty arises from the second term, and we can show, as in [7], that exists a function $C(\delta) \underset{\delta\to 0}{\to} 0$ such that

$$B_2^N \le \exp\{C(\delta)N\}.$$

On the other hand,

$$B_1^N \overset{\text{Hölder}}{\le} \left(\int\mathcal{E}_\nu\left[\exp\left\{\rho^2\int_0^T\big(G_t(x)+m_\nu(t,x)\big)dW_t(x,r) - \frac{\rho^4}{2}\int_0^T\big(G_t(x)+m_\nu(t,x)\big)^2dt\right\}\right]dP(x,r)\right)^{\frac{N}{\rho}}$$
$$\times\left(\int\mathcal{E}_\nu\left[\exp\left\{\frac{(\rho+1)\rho^2}{2}\int_0^T\big(G_t(x)+m_\nu(t,x)\big)^2dt\right\}\right]dP(x,r)\right)^{\frac{N(\rho-1)}{\rho}}$$

The first term in the righthand side equals one by Fubini Theorem and martingale property. Moreover, under the hypothesis $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$ and for $\rho-1$ small enough, we can show using (12), that exists a constant C_T, uniform on $\mathcal{C}$, such that

$$\mathcal{E}_\nu\left[\exp\left\{\frac{(\rho+1)\rho^2}{2}\int_0^T\big(G_t(x)+m_\nu(t,x)\big)^2dt\right\}\right] \le C_T.$$

Hence, if we take ρ close enough to 1, and under the hypothesis $\frac{2\sigma^2\|b\|_\infty^2 T}{\lambda^2} < 1$, we can find a finite constant $C_2(\rho), \lim_{\rho\to 1} C_2(\rho) = 0$, such that:

$$B_1^N \leq e^{C_2(\rho)N} \tag{24}$$

□

References

[1] G Ben Arous, Amir Dembo, and Alice Guionnet. Aging of spherical spin glasses. *Probability theory and related fields*, 120(1):1–67, 2001.

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

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

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

[5] L Boltzmann. *Lectures on gas theory*. Dover, New York, 1895.

[6] Sem Borst and Florian Simatos. A stochastic network with mobile users in heavy traffic. *Queueing Systems*, 74(1):1–40, 2013.

[7] Tanguy Cabana and Jonathan Touboul. Large deviations and dynamics of large stochastic heterogeneous neural networks. *Journal of Statistical Physics*, 2013.

[8] Tanguy Cabana and Jonathan Touboul. Large deviations for spatially extended random neural networks. *(submitted)*, 2015.

[9] José A Carrillo, Massimo Fornasier, Giuseppe Toscani, and Francesco Vecil. Particle, kinetic, and hydrodynamic models of swarming. In *Mathematical modeling of collective behavior in socio-economic and life sciences*, pages 297–336. Springer, 2010.

[10] Hiroaki Daido. Quasientrainment and slow relaxation in a population of oscillators with random and frustrated interactions. *Physical review letters*, 68(7):1073, 1992.

[11] Hiroaki Daido. Algebraic relaxation of an order parameter in randomly coupled limit-cycle oscillators. *Physical Review E*, 61(2):2145, 2000.

[12] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Verlag, 2009.

[13] Frank Den Hollander. *Large deviations*, volume 14. American Mathematical Soc., 2008.

[14] A Destexhe, ZF Mainen, and TJ Sejnowski. An efficient method for computing synaptic conductances based on a kinetic model of receptor binding. *Neural Computation*, 6(1):14–18

[15] J.D. Deuschel and D.W. Stroock. *Large deviations*, volume 137. Academic Pr, 1989.

[16] Olivier Faugeras and James Maclaurin. Asymptotic description of stochastic neural networks. i. existence of a large deviation principleastic neural networks. ii. characterization of the limit law. *Comptes Rendus Mathematique*, 352(10):847–852, 2014.

[17] Carl Graham, Philippe Robert, et al. Interacting multi-class transmissions in large stochastic networks. *The Annals of Applied Probability*, 19(6):2334–2361, 2009.

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

[19] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500–544, 1952.

[20] Yoshiki Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International symposium on mathematical problems in theoretical physics*, pages 420–422. Springer, 1975.

[21] Eric Luçon, Wilhelm Stannat, et al. Mean field limit for disordered diffusions with singular interactions. *The Annals of Applied Probability*, 24(5):1946–1993, 2014.

[22] Sylvie Méléard. Asymptotic behaviour of some interacting particle systems; mckean-vlasov and boltzmann models. In *Probabilistic models for nonlinear partial differential equations*, pages 42–95. Springer, 1996.

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

[24] Lorenzo Pareschi and Giuseppe Toscani. *Interacting multiagent systems: kinetic equations and Monte Carlo methods*. Oxford University Press, 2013.

[25] D. Sherrington and S. Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792–1796

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

[27] JC Stiller and G Radons. Dynamics of nonlinear oscillators with random interactions. *Physical Review E*, 58(2):1789, 1998.

[28] JC Stiller and G Radons. Self-averaging of an order parameter in randomly coupled limit-cycle oscillators. *Physical Review E*, 61(2):2148, 2000.

[29] AS Sznitman. Equations de type de boltzmann, spatialement homogenes. *Probability Theory and Related Fields*, 66(4):559–592, 1984.

[30] AS Sznitman. Topics in propagation of chaos. *Ecole d'Eté de Probabilités de Saint-Flour XIX*, pages 165–251, 1989.