Annealed and Quenched Propagation of Chaos for Asymmetric Spin Glass Dynamics

G. BEN AROUS
A. GUIONNET

LMENS - 95 - 5
Annealed and Quenched Propagation of Chaos for Asymmetric Spin Glass Dynamics

G. BEN AROUS
A. GUIONNET

LMENS - 95 - 5

March 1995

Laboratoire de Mathématiques de l’Ecole Normale Supérieure
45 rue d’Ulm 75230 PARIS Cedex 05
Tel : (33)(1) 44 32 00 00
Adresse électronique : .. @dmi.ens.fr

*Laboratoire de Modélisation Stochastique et Statistique, URA 743
Université Paris-Sud (Bât 425) 91405 ORSAY Cedex
Annealed and Quenched Propagation of Chaos for
Asymmetric Spin Glass Dynamics

G. Ben Arous*, A. Guionnet†

Abstract: We study the asymptotic behavior of asymmetrical spin glass dynamics in a Sherrington-Kirkpatrick model as proposed by Sompolinsky-Zippelius. We prove, without any condition on time and temperature, an annealed propagation of chaos result. Extending this result to replicated systems, we conclude that the law of a single spin converges to a non markovian probability measure, in law with respect to the random interaction.

Keywords: Large deviations, Interacting random processes, Statistical mechanics, Langevin dynamics.

code A.M.S: 60F10, 60H10, 60K35, 82C44, 82C31, 82C22

*URA 762, CNRS, DMI, Ecole Normale Superieure, 75230 Paris, France
†URA 743, CNRS, Bat. 425, Université de Paris Sud, 91405 Orsay, France
1 Introduction

The goal of this paper is to prove an annealed and a quenched propagation of chaos result with no restriction on time or temperature for the asymmetric Langevin spin glass dynamics as introduced by Sompolinski and Zippelius [8] and that we have studied in [2] in a short time and high temperature regime.

More precisely, we study the following system \( S_N^N (J) \) of randomly interacting diffusions:

\[
S_N^N (J) = \left\{ \begin{array}{ll}
\dot{x}_i^j = -\nabla U (x_i^j) + dB_i^j + \frac{\beta}{\sqrt{N}} \sum_{i=1}^{N} J_{ji} x_i^j dt & \forall 1 \leq i \leq N \\
Law of x_0 = \mu_0^N
\end{array} \right.
\]

where \((B^j)_{1 \leq j \leq N}\) is a \(N\) dimensional Brownian motion.

As in [2], we will study only bounded spins, i.e we will assume that \(\mu_0\) is a probability measure on a bounded interval \([-A, A]\) which does not put mass on the boundary \([-A, +A]\) and that \(U(x)\) is defined on \([-A, A]\) and tends to infinity when \(|x| \to A\) sufficiently fast to insure our spins \(x^i\) stay in the interval \([-A, A]\).

Moreover, we will study asymmetric dynamics, i.e we will assume that the whole matrix \((J_{ij})_{i,j}\) is made of i.i.d \(N(0,1)\) random variables and we will not impose the symmetry \(J_{ij} = J_{ji}\).

We prove that the annealed law of the empirical measure \(\hat{\mu}^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x^i}\) converges to \(\delta_Q\) where \(Q\) is the non markovian law described in [2], with no restriction on time and temperature (see section 2).

We deduce from this an annealed propagation of chaos, that is convergence of the annealed law of \((x^1, \cdot \cdot \cdot, x^m)\) to \(Q^\otimes m\) (see theorem 2.2).

We have seen in [2], section 6, that \(Q\) can be seen as an average of non Markovian processes, namely that there exists a centered gaussian process \(h\) and a non markovian probability measure \(P_h\) such that \(Q = \mathcal{E}^h [P_h]\), where \(\mathcal{E}^h\) denotes the expectation over \(h\). So that, if we denote \(\mathcal{E}\) the expectation over the random couplings \(J\) and \(P_{J^N}^N\) the weak solution of \(S^N_J\) until a time \(T\), we have proved that, for any bounded continuous function \(f\), for any temperature \(\frac{1}{\beta}\):

\[
\lim_{N \to \infty} \mathcal{E} \left[ \int f(x^1) dP_{J^N}^N \right] = \mathcal{E}^h \left[ \int f dP_h \right]
\]

Using replica, we improve this last result and deduce a quenched propagation of chaos, i.e prove that \(\int f(x^1) dP_{J^N}^N\) converges in law to \(\int f(x) dP_h\) (see section 3).

We thank Stefano Olla for many stimulating discussions and, in particular, for the crucial suggestion to prove the tightness result of section 5 by entropy estimates.
2 Annealed propagation of chaos

The aim of this section is to prove:

**Theorem 2.1** The law of the empirical measure \( \hat{\mu}^N = \frac{1}{N} \sum_{i=1}^N \delta_{\omega^i} \) under \( Q_\beta^N = \mathcal{E} \left[ P_\beta^N (J) \right] \) converges to \( \delta_Q \).

Where \( Q \) is described in theorem 2.5.

According to lemma 3.1 of Szmitman [7], theorem 2.1 gives the annealed propagation of chaos result:

**Theorem 2.2** \( Q_\beta^N = \mathcal{E} \left[ P_\beta^N (J) \right] \) is \( Q \)-chaotic, i.e for any bounded continuous functions \( (f_1, \cdots, f_m) \),

\[
\lim_{N \to \infty} \frac{1}{N} \log Q_\beta^N (\hat{\mu}^N \in K) \leq - \inf_K H.
\]

In [2], we proved theorem 2.1 in the high temperature and small time regime \( \beta^2 A^2 T < 1 \). This last restriction was mainly due to an exponential tightness criterion which was crucial to prove that the annealed law of the empirical measure satisfies a large deviation principle which gave theorem 2.1. Here, we deduce theorem 2.1 from a weak large deviation principle and from a tightness result. Namely, if \( W^A_T \) denotes the set of continuous functions from \([0, T]\) into \([-A, +A]\) and \( \mathcal{M}_1^+ (W^A_T) \) the set of probability measures on \( W^A_T \), we prove that:

**Theorem 2.3** There exists a good rate function \( H \) such that, for any compact subset \( K \) of \( \mathcal{M}_1^+ (W^A_T) \):

\[
\lim_{N \to \infty} \frac{1}{N} \log Q_\beta^N (\hat{\mu}^N \in K) \leq - \inf_K H.
\]

And:

**Theorem 2.4** For any real number \( \epsilon > 0 \), there exists a compact subset \( K_\epsilon \) of \( \mathcal{M}_1^+ (W^A_T) \) such that, for any integer number \( N \),

\[
Q_\beta^N (\hat{\mu}^N \in K_\epsilon) \leq \epsilon.
\]

Theorem 2.3 is proved in section 4 and theorem 2.4 is proved in section 5.

To deduce theorem 2.1 from theorems 2.3 and 2.4, we need to recall that we characterized the minima of \( H \) in [2], section 5, and proved that:

**Theorem 2.5**

1) \( H \) achieves its minimal value at the probability measures \( Q \) which satisfy:

\[ Q \ll P \quad \frac{dQ}{dP} = \mathcal{E} \left[ \exp \{ \beta \int_0^T G_t^Q dB_t - \frac{\beta^2}{2} \int_0^T (G_t^Q)^2 dt \} \right] \tag{1} \]

Where \( \mathcal{E} \) denotes the expectation over a centered gaussian process \( G_t^Q \) with covariance

\[ \mathcal{E} \left[ G_t^Q G_s^Q \right] = \int x_t x_s dQ (x) \]

2) There exists a unique probability measure \( Q \) which satisfies (1).
Proof of Theorem 2.1

Let $\delta$ be a strictly positive real number and denote $B(Q, \delta)$ the open ball with respect to a metric compatible with the weak topology on $\mathcal{M}_1^+(W^A_T)$, for instance the Vaserstein metric (which definition is given in (6)). We prove that $Q^N_\delta(\hat{\mu}^N \in B(Q, \delta)^c)$ converges to zero for any positive real number $\delta$. Indeed, if $K_\varepsilon$ are the compact sets defined in theorem 2.4, we have, for any $\varepsilon > 0$:

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

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

(2)

But, since $B(Q, \delta)^c$ is closed, $K_\varepsilon \cap B(Q, \delta)^c$ is compact so that we can use theorem 2.3:

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

But $\inf_{K_\varepsilon \cap B(Q, \delta)^c} H$ is strictly positive according to theorem 2.5. Hence, (2) implies that, for any $\varepsilon > 0$,

$$\lim_{N \to \infty} Q^N_\delta(\hat{\mu}^N \in B(Q, \delta)^c) \leq \varepsilon$$

i.e

$$\lim_{N \to \infty} Q^N_\delta(\hat{\mu}^N \in B(Q, \delta)^c) = 0$$

3 Quenched propagation of chaos and Replica

Theorem 2.2 can be extended to replicated systems as follows:

Let $r$ be an integer number and denote $Q^r_\beta$ the annealed law of replicated spin glass dynamics:

$$Q^r_\beta = \mathbb{E}[P^N_\beta(J)^\otimes r]$$

Let $Q_\varepsilon$ be defined by:

$$Q_\varepsilon \ll P_\varepsilon \quad \frac{dQ_\varepsilon}{dP_\varepsilon} = \mathbb{E}\left[\exp\{\beta \int_0^T < G_t^Q, dB_t > - \frac{\beta^2}{2} \int_0^T ||G_t^Q||^2 dt\}\right]$$

(3)

Where $G^Q$ is a $r$-dimensional centered gaussian process with covariance:

$$\mathbb{E}[G^Q_t x_i G^Q_s x_j] = \int x_s x_i dQ_r(x)$$

And where $\langle \ , \ \rangle$ denotes the euclidean scalar product in $\mathbb{R}^r$ and $||\ ||^2 = \langle \ , \ \rangle$.

Then $Q_\varepsilon$ exists and is unique (see [2], section 6), and we have:
Theorem 3.1

For any integer number \( r \), the law of the empirical measure \( \hat{\mu}_r^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i^r} \) under \( Q_r^N \) converges to \( \delta_{Q_r} \).

The proof of theorem 3.1 is very similar to that of theorem 2.2. We omit it.

As a consequence :

**Theorem 3.2** \( Q_r^N \) is \( Q_r \)-chaotic, i.e for any bounded continuous functions \( (F_1, \cdots, F_m) \) on \((W_T^A)^r\),

\[
\lim_{N \to \infty} \mathcal{E} \left[ \prod_{i=1}^{m} F_i(x_1^r, \cdots, x_r^r) dP^N (J) \right] = \prod_{i=1}^{m} \int F_i(x_1^r, \cdots, x_r^r) dQ_r
\]

In particular, for any bounded continuous functions \( (f_1, \cdots, f_m) \) on \( W_T^A \),

\[
\lim_{N \to \infty} \mathcal{E} \left[ \left( \int \prod_{i=1}^{m} f_i(x_i^r) dP^N (J) (x) \right) \right] = \prod_{i=1}^{m} \int f_i(x_1^r) \cdots f_i(x_r^r) dQ_r
\]

To deduce a quenched propagation of chaos from theorem 3.2, we need to identify the probability measures \( Q_r \) themselves as replicated laws. This was done in [2], section 6, where we proved that there exists a couple \( (h, P_h) \) of a gaussian process \( h \) and a probability measure \( P_h \) on \( W_T^A \) (which depends on \( h \)) such that :

**Theorem 3.3**

For any integer \( r \)

\[
Q_r = \mathcal{E}^h \left[ P_h^{\otimes r} \right]
\]

Moreover, the couple \( (h, P_h) \) is defined by the following non linear procedure :

For \( f \) in \( L^2([0, T]) \), let \( P(f) \) be the restriction on \([0, T]\) of the law of the diffusion

\[
\begin{cases}
    \frac{dx_i}{dt} = -\nabla U(x_i) dt + dB_i + \beta f(t) dt \\
    \text{Law of } x_0 = \mu_0
\end{cases}
\]

Let \( (h, g) \) be two independent centered gaussian processes and denote \( \mathcal{E}^h \) (resp. \( \mathcal{E}^g \)) the expectation over \( h \) (resp. \( g \)). We define non linearly the covariances of \( (h, g) \) by :

\[
\mathcal{E}^g[g,g] = \mathcal{E}^h \mathcal{E}^g \left[ \int x_s x_t dP(g + h) \right] - \mathcal{E}^h \left[ \mathcal{E}^g \left[ \int x_s dP(g + h) \right] \mathcal{E}^g \left[ \int x_t dP(g + h) \right] \right]
\]

\[
\mathcal{E}^h[h,h] = \mathcal{E}^h \left[ \mathcal{E}^g \left[ \int x_s dP(g + h) \right] \mathcal{E}^g \left[ \int x_t dP(g + h) \right] \right]
\]

Finally \( P_h \) is given by :

\[
P_h = \mathcal{E}^g \left[ P(g + h) \right]
\] (4)

Theorem 3.2 and theorem 3.3 give :
Corollary 3.4 For any integer $r$, for any bounded continuous functions $(f_1, \ldots, f_m)$ on $W_T^A$,
\[
\lim_{N \to \infty} \mathcal{E} \left[ \left( \int f_1(x^1) \cdots f_m(x^m) dP^N_{\beta}(J)(x) \right)^r \right] = \prod_{i=1}^{m} \mathcal{E}^h \left[ \left( \int f_i(x) dP_h(x) \right)^r \right]
\]
Since the random variables $\int f_1(x^1) \cdots f_m(x^m) dP^N_{\beta}(J)(x)$ are bounded, corollary 3.4 is equivalent to the convergence in law of such random variables, which gives the quenched propagation of chaos result:

Theorem 3.5
For any bounded continuous functions $(f_1, \ldots, f_m)$ on $W_T^A$, $\int f_1(x^1) \cdots f_m(x^m) dP^N_{\beta}(J)(x)$ converges in law, when $N$ tends to infinity, to $\prod_{j=1}^{m} \int f_j dP_{h_j}$, where $h_j$ are independent copies of the centered gaussian process $h$ described above.

Moreover, we described in [2], section 6, the case where the limiting law is deterministic, i.e the case where $h$ is null almost surely. Then $P_h = \mathcal{E}^h [P_h] = Q$. Roughly speaking, it is the case where the potential $U$ is even and the initial law is symmetric. Then, theorem 3.5 becomes:

Corollary 3.6
If $U$ is even and $\mu_0$ is symmetric, for any bounded continuous functions $(f_j, 1 \leq j \leq m)$, $\int f_1(x^1) \cdots f_m(x^m) dP^N_{\beta}(J)$ converges in probability to $\prod_{j=1}^{m} \int f_j dQ$.

Remark 3.7:
1) Theorem 3.5 (and corollary 3.6) can also be stated for finite vectors $\left( \int f_i^1(x^1) \cdots f_i^m(x^m) dP^N_{\beta}(J) \right)_{1 \leq i \leq n}$, where $\left( f_i^j, 1 \leq i \leq n, 1 \leq j \leq m \right)$ are bounded continuous functions, and one finds that $\left( \int f_i^1(x^1) \cdots f_i^m(x^m) dP^N_{\beta}(J) \right)_{1 \leq i \leq n}$ converges in law to $\left( \prod_{j=1}^{m} \int f_j^j dP_{h_j} \right)_{1 \leq i \leq n}$. This is obvious since theorem 3.2 gives the convergence of every moments of these random variables.

2) Theorem 3.2 gives a complete description of the microscopic behaviour of asymmetric spin glass dynamics: it shows that the law of a finite number of spins converges to independent laws submitted to independent identically distributed gaussian external fields. Moreover, it is clear that the limiting law $P_h$ is non markovian since it depends on all the past through the law of the gaussian process $g$ (see (4)).
4 Annealed weak large deviation upper bound

To state precisely the main result of this section, i.e. the weak large deviation upper bound for the annealed law of the empirical measure stated in theorem 2.3, we first define the rate function $H$ which governs these deviations.

**Definition 4.1** For any $\mu \in \mathcal{M}_1^+(W_T^A)$, we define a centered gaussian process $G^\mu$ by its covariance:

$$\mathbb{E}[G_s^\mu G_t^\mu] = \int x_s x_t d\mu(x)$$

(5)

For any probability measure $\mu$ which is absolutely continuous with respect to $P$, we then define $\Gamma(\mu)$ by:

$$\Gamma(\mu) = \int \log \mathbb{E} \left[ \exp \left\{ \beta \int_0^T G_t^\mu dB_t(x) - \frac{\beta^2}{2} \int_0^T (G_t^\mu)^2 dt \right\} \right] d\mu(x)$$

Where $B_t(x) = x_t - x_0 + \int_0^t \nabla U(x_s) ds$ for any $x \in W_T^A$.

Let then $I(\mu | P)$ denotes the entropy of $\mu$ with respect to $P$, namely:

$$I(\mu | P) = \left\{ \begin{array}{ll} \int \log \frac{d\mu}{dP} d\mu & \text{if } \mu << P \\ +\infty & \text{otherwise} \end{array} \right.$$

Then, we define a function $H$ on $\mathcal{M}_1^+(W_T^A)$ by:

$$H(\mu) = \left\{ \begin{array}{ll} I(\mu | P) - \Gamma(\mu) & \text{if } I(\mu | P) < \infty \\ +\infty & \text{otherwise} \end{array} \right.$$

**Remark 4.2:**

1) One can construct the gaussian processes $\{G^\mu, \mu \in W_T^A\}$ on the same probability space. For instance, let $(\Omega, \gamma)$ be a probability space and $(J_i)_{i \in \mathbb{N}}$ be i.i.d $N(0,1)$ on $\Omega$, then, if $(e_i^\mu)_{i \in \mathbb{N}}$ is an orthonormal basis in $L_2^\mu(W_T^A)$,

$$G_s^\mu = \sum_{i \in \mathbb{N}} J_i \int x_s e_i^\mu(x) d\mu(x)$$

is a centered gaussian process with covariance (5). In the following pages, we will suppose that we have constructed all the gaussian processes $\{G^\mu, \mu \in W_T^A\}$ on the same probability space $(\Omega, \gamma)$ and denote $\mathbb{E}$ the expectation under $\gamma$.

2) For any $\mu \ll P$, $\Gamma(\mu)$ is well defined since $B$ is a semimartingale. As a consequence, $\Gamma$ is well defined on $\{ \mu \in \mathcal{M}_1^+(W_T^A) / I(\mu | P) < \infty \}$, so that $H$ is well defined.

Then:

**Theorem 4.3**

1) $H$ is a good rate function, i.e $H$ is positive and, for any real number $M$, $\{ \mu \in \mathcal{M}_1^+(W_T^A) / H(\mu) \leq M \}$ is compact.
2) For any compact subset $K$ of $\mathcal{M}_1^+(W^A_T)$:

$$\lim_{N \to \infty} \frac{1}{N} \log Q^N_{\beta} \left( \hat{\mu}^N \in K \right) \leq - \inf_K H.$$

**Proof of theorem 4.3.**

The first point is proved in [2], section 4.

To prove the second point, we first notice that, according to lemma 3.6 of [2], we have:

$$dQ^N_{\beta} = \exp \{ N \Gamma (\hat{\mu}^N) \} \, dP^\otimes N.$$

Thus, if $\Gamma$ was bounded and continuous, theorem 4.3.2) would be clear. To circumvent the fact that none of these properties is satisfied, we shall approximate $\Gamma$ by linear functions. More precisely, for any $\nu \in \mathcal{M}_1^+(W^A_T)$, we define a map $\Gamma_{\nu}$ from $\mathcal{M}_1^+(W^A_T)$ into $\mathbb{R}$ by:

$$\Gamma_{\nu}(\mu) = \int \log \mathcal{E} \left[ \exp \left\{ \beta \int_0^T G^\nu_s dB_s(x) \right. \right.$$

$$- \frac{\beta^2}{2} \int_0^T (G^\nu_s)^2 ds \left. \right\} \right] d\mu(x).$$

We denote $d$ the Vaserstein distance on $W^A_T$, i.e.:

$$d(\mu, \nu) = \sup \{ \int \sup_{s \leq T} |x_s - y_s| d\xi(x, y) \} \quad (6)$$

Where the supremum is taken on the probability measure $\xi$ with marginals $\mu$ and $\nu$.

The key of our proof is the following technical lemma:

**Lemma 4.4** For any real number $\alpha, \alpha > 1$, there exists a strictly positive real number $\delta_\alpha$ such that, for any $\delta < \delta_\alpha$, there exists a function $C_{\alpha}(\cdot)$ in $\mathbb{R}$ such that $\lim_{\delta \to 0} C_{\alpha}(\delta) = 0$ and:

$$\int_{d(\hat{\mu}^N, \nu) \leq \delta} \exp \left\{ \alpha N (\Gamma - \Gamma_{\nu})(\hat{\mu}^N) + N \Gamma_{\nu}(\hat{\mu}^N) \right\} dP^\otimes N \leq \exp C_{\alpha}(\delta) N. \quad (7)$$

**Proof of lemma 4.4.**

Let $B^N = \int_{d(\hat{\mu}^N, \nu) \leq \delta} \exp \left\{ \alpha N (\Gamma - \Gamma_{\nu})(\hat{\mu}^N) \right\} \exp \{ N \Gamma_{\nu}(\hat{\mu}^N) \} dP^\otimes N$

Let $Q_{\nu}$ be the probability measure on $W^A_T$ defined by:

$$dQ_{\nu} = \mathcal{E} \left[ \exp \left\{ \beta \int_0^T G^\nu_s dB_s - \frac{\beta^2}{2} \int_0^T (G^\nu_s)^2 ds \right\} \right] dP$$

Then:

$$d(Q_{\nu})^\otimes N = \exp \{ N \Gamma_{\nu}(\hat{\mu}^N) \} dP^\otimes N$$

Writing down the definitions of $\Gamma$ and $\Gamma_{\nu}$, we find:

8
\[ B^N = \int_{d(\hat{\nu}, \nu) < \delta} \prod_{j=1}^{N} \left( \frac{\mathcal{E} \left[ \exp \{ \beta \int_0^T G^{\hat{\nu}}_t dB^j_t - \frac{\beta^2}{2} \int_0^T (G^{\hat{\nu}}_t)^2 dt \} \right]}{\mathcal{E} \left[ \exp \{ \beta \int_0^T G^{\nu}_t dB^j_t - \frac{\beta^2}{2} \int_0^T (G^{\nu}_t)^2 dt \} \right]} \right)^{\alpha} d(\nu) \circ t \]

Assume that we have constructed on the same probability space \((\Omega, \gamma)\) the gaussian processes \(G^{\hat{\nu}}\) and \(G^{\nu}\) and denote:

\[ X^j_N = \beta \int_0^T G^{\hat{\nu}}_s dB^j_s - \frac{\beta^2}{2} \int_0^T (G^{\hat{\nu}}_s)^2 ds \]
\[ X^j_\nu = \beta \int_0^T G_s dB^j_s - \frac{\beta^2}{2} \int_0^T (G_s)^2 ds \]

So that:

\[ B^N = \int_{d(\hat{\nu}, \nu) < \delta} \prod_{j=1}^{N} \left( \frac{\mathcal{E} \left[ \exp \{ X^j_N \} \right]}{\mathcal{E} \left[ \exp \{ X^j_\nu \} \right]} \right)^{\alpha} d(\nu) \circ t \]

\[ = \int_{d(\hat{\nu}, \nu) < \delta} \prod_{j=1}^{N} \left( \mathcal{E} \left[ \exp \frac{X^j_\nu}{\exp X^j_\nu} \exp \left( X^j_N - X^j_\nu \right) \right] \right)^{\alpha} d(\nu) \circ t \]

Since \(\alpha > 1\), we can apply Jensen inequality in the last equality:

\[ B^N \leq \int_{d(\hat{\nu}, \nu) < \delta} \prod_{j=1}^{N} \mathcal{E} \left[ \exp \frac{X^j_\nu}{\exp X^j_\nu} \exp \alpha \left( X^j_N - X^j_\nu \right) \right] d(\nu) \circ t \]

Therefore, if \(p^{-1} + q^{-1} = 1\),

\[ B^N \leq \left\{ \prod_{j=1}^{N} \frac{\mathcal{E} \left[ \exp p X^j_\nu \right]}{\mathcal{E} \left[ \exp X^j_\nu \right]} \right\}^{\frac{1}{p}} \times \left\{ \int_{d(\hat{\nu}, \nu) < \delta} \prod_{j=1}^{N} \mathcal{E} \left[ \exp \alpha q \left( X^j_N - X^j_\nu \right) \right] d(\nu) \circ t \right\}^{\frac{1}{q}} \]

(9)

We first bound the first term in the right hand side of (9). Gaussian calculus shows that

\[ \exists C < \infty \quad \frac{\mathcal{E} \left[ \exp \{ p X^j_\nu \} \right]}{\left( \mathcal{E} \left[ \exp X^j_\nu \right] \right)^p} \leq \exp \left\{ C(p - 1) \left( 1 + \mathcal{E} \left[ \left( \int_0^T G_t dB^j_t \right)^2 \right] \right) \right\} \]

So that, if \(p\) is close enough to 1, we can find a finite constant \(C(p), C(p) \searrow 0\) when \(p \searrow 1\), such that:

\[ B_1^N = \int_{d(\hat{\nu}, \nu) < \delta} \prod_{j=1}^{N} \left( \frac{\mathcal{E} \left[ \exp p X^j_\nu \right]}{\mathcal{E} \left[ \exp X^j_\nu \right]} \right)^{\alpha} d(\nu) \circ t \]

\[ = \left( \int \frac{\mathcal{E} \left[ \exp p X^j_\nu \right]}{\left( \mathcal{E} \left[ \exp X^j_\nu \right] \right)^p} d(\nu) \right)^N \leq e^{C(p)N} \]

(10)
We now bound the second term in the right hand side of (9) using H"older inequality with conjugate exponents $(\eta, \sigma)$ ($\eta^{-1} + \sigma^{-1} = 1$):

\[
B_2^N = \int_{d(\tilde{\mu}^N, \nu) < \delta} \left( \prod_{j=1}^N \mathcal{E} \left[ \exp \alpha q \left( X^j - X^j_{\nu} \right) \right] \right) \exp N \Gamma_\nu(\tilde{\mu}^N) dP^{\otimes N}(x)
\]

\[
\leq \left\{ \int \exp \{ N \sigma \Gamma_\nu(\tilde{\mu}^N) \} dP^{\otimes N} \right\}^{\frac{1}{\sigma}}
\times \left\{ \int_{d(\tilde{\mu}^N, \nu) < \delta} \left( \prod_{j=1}^N \mathcal{E} \left[ \exp \{ \alpha \eta q \left( X^j - X^j_{\nu} \right) \} \right] \right) dP^{\otimes N} \right\}^{\frac{1}{\eta}}
\]

The first term in the right hand side of (11) can be bounded if $\sigma - 1$ is small enough:

\[
\int \exp \{ N \sigma \Gamma_\nu(\tilde{\mu}^N) \} dP^{\otimes N}(x) = \left( \int \mathcal{E} \left[ \exp \left\{ \beta \int_0^T G_t^\nu dB_t(x) - \frac{\beta^2}{2} \int_0^T (G_t^\nu)^2 dt \right\} \right] dP(x) \right)^N
\]

\[
\leq \left( \mathcal{E} \left[ \int \exp \{ \beta \sigma \int_0^T G_t^\nu dB_t(x) - \frac{\beta^2}{2} \sigma \int_0^T (G_t^\nu)^2 dt \} dP(x) \right] \right)^N
\]

\[
= \left( \mathcal{E} \left[ \int \exp \left\{ \frac{\beta^2}{2} (\sigma^2 - \sigma) \int_0^T (G_t^\nu)^2 dt \right\} \right] \right)^N
\]

But gaussian integrability properties imply, as detailed in appendix A of [2], that there exists a finite constant $C(\sigma)$, $\lim_{\sigma \to 1} C(\sigma) = 0$, such that:

\[
\mathcal{E} \left[ \int \exp \left\{ \frac{\beta^2}{2} (\sigma^2 - \sigma) \int_0^T (G_t^\nu)^2 dt \right\} \right] \leq \exp \{ C(\sigma) \}
\]

(12)

So that we have proved that:

\[
\int \exp \{ N \sigma \Gamma_\nu(\tilde{\mu}^N) \} dP^{\otimes N}(x) = \left( \int \exp \{ \sigma \Gamma_\nu(\tilde{\mu}^N) \} dP(x) \right)^N < e^{C(\sigma)N}
\]

(13)

We bound the second term in the right hand side of (11). By Cauchy Schwarz inequality, if $\kappa = \alpha \eta q \beta$:

\[
\int_{d(\tilde{\mu}^N, \nu) < \delta} \left( \prod_{j=1}^N \mathcal{E} \left[ \exp \alpha \eta q \left( X^j - X^j_{\nu} \right) \right] \right) dP^{\otimes N}
\]

\[
\leq \left\{ \int \prod_{j=1}^N \mathcal{E} \left[ \exp \{ 2 \kappa \int_0^T (G_s^N - G_s^\nu) dB_s^j - 2 \kappa^2 \int_0^T (G_s^N - G_s^\nu)^2 ds \} \right] dP^{\otimes N} \right\}^{\frac{1}{2}}
\times \left\{ \int \prod_{j=1}^N \mathcal{E} \left[ \exp \{ 2 \kappa^2 \int_0^T (G_s^N - G_s^\nu)^2 ds + \beta \kappa \int_0^T (G_s^N - G_s^\nu)^2 ds \} \right] dP^{\otimes N} \right\}^{\frac{1}{2}}
\]

(14)
The first term is bounded by one by supermartingale properties. For the second term, we remark that:

\[ \int_0^T \left( (G_s^j)^2 - (G_s)^2 \right) ds \leq \frac{1}{2} \frac{1}{\delta} \left( \frac{1}{\delta} \int_0^T (G_s^j - G_s)^2 ds + \int_0^T (G_s^j + G_s)^2 ds \right) \]

(15)

Hence, we can apply as in (12) lemma A.3.2 of [2] to the right hand side of (15) so that we find that, for any real number \( \kappa \), there exists \( C_\kappa(\delta) \), \( \lim_{\delta \to 0} C_\kappa(\delta) = 0 \), such that, for any \( (x^i)_{1 \leq i \leq N} \) such that \( d\left( \tilde{\mu}^N, \nu \right) < \delta \):

\[ \mathcal{E} \left[ \exp \left\{ 2\kappa^2 \int_0^T (G_s^N - G_s^\nu)^2 ds + \beta \kappa \int_0^T \left( (G_s^N)^2 - (G_s^\nu)^2 \right) ds \right\} \right] \leq e^{C_\kappa(\delta)} \]

So that:

\[ \int_{d(\tilde{\mu}^N, \nu) < \delta} \mathcal{E} \left[ \exp \left\{ 2\kappa^2 \int_0^T (G_s^N - G_s^\nu)^2 ds + \beta \kappa \int_0^T \left( (G_s^N)^2 - (G_s^\nu)^2 \right) ds \right\} \right]^N dP^{\otimes N} \leq e^{C_\kappa(\delta)N} \]

(16)

Therefore, inequalities (9), (10), (11), (13) and (16) show that we can choose \( p \) close enough to 1 and then \( \delta \) small enough so that there exists a finite real number \( C_\alpha(\delta) \), \( \lim_{\delta \to 0} C_\alpha(\delta) = 0 \), such that:

\[ \frac{1}{N} \log \int_{d(\tilde{\mu}^N, \nu) < \delta} \exp \left\{ \alpha N (\Gamma - \Gamma^\nu) \left( \tilde{\mu}^N \right) \right\} \exp N \Gamma^\nu \left( \tilde{\mu}^N \right) dP^{\otimes N} < \exp C_\alpha(\delta)N \]

We finally prove the weak large deviation upper bound theorem 4.3.2):

Let \( K \) be a compact set of \( \mathcal{M}^+_1(W_T^A) \), \( K \) can be covered by a finite union of open balls for the Vaserstein’s metric:

\[ K \subseteq \bigcup_{i=1}^M B(\nu_i, \delta) \]

Where

\[ B(\nu_i, \delta) = \left\{ \mu \in \mathcal{M}^+_1(W_T^A) / d(\mu, \nu_i) < \delta \right\} \]

According to (23), we have:

\[ Q_\beta^N(\tilde{\mu}^N \in K) = \int_K \exp \left\{ N \Gamma(\tilde{\mu}^N) \right\} dP^{\otimes N} \]

\[ = \sum_{i=1}^M \int_{K \cap B(\nu_i, \delta)} \exp N \left\{ \Gamma(\tilde{\mu}^N) - \Gamma_{\nu_i}(\tilde{\mu}^N) \right\} \exp \left\{ N \Gamma_{\nu_i}(\tilde{\mu}^N) \right\} dP^{\otimes N} \]

With the definition of the probability measures \( Q_\nu \) as in (8), we get:

\[ Q_\beta^N(\tilde{\mu}^N \in K) = \sum_{i=1}^M \int_{K \cap B(\nu_i, \delta)} \exp N \left\{ \Gamma(\tilde{\mu}^N) - \Gamma_{\nu_i}(\tilde{\mu}^N) \right\} dQ_{\nu_i}^{\otimes N} \]
Hölder inequality shows that, for \( p, q \) conjugate exponents:

\[
Q^N_{\beta} (\hat{\mu}^N \in K) \leq \sum_{i=1}^{M} \left( \int_{d(\hat{\mu}^N, \nu_i) < \delta} \exp qN\{\Gamma(\hat{\mu}^N) - \Gamma_{\nu_i}(\hat{\mu}^N)\} dQ_{\nu_i}^N \right)^{\frac{1}{q}} Q_{\nu_i}^\cap N (B(\nu_i, \delta) \cap K)^{\frac{1}{p}}
\]

So that proposition 4.4 implies that:

\[
Q^N_{\beta} (\hat{\mu}^N \in K) \leq \exp \frac{1}{q} NC_q(\delta) \times \left\{ \sum_{i=1}^{M} Q_{\nu_i}^\cap N (B(\nu_i, \delta) \cap K)^{\frac{1}{p}} \right\}
\]

Thus Sanov Theorem implies that:

\[
\lim_{N \to \infty} \frac{1}{N} \log Q^N_{\beta} (\hat{\mu}^N \in K) \leq - \inf p \inf_{1 \leq i \leq M} \inf_{B(\nu_i, \delta) \cap K} I(\mid | Q_{\nu_i} \mid) + \frac{1}{q} C_q(\delta)
\]  \( (17) \)

But one can see (as proved in [2], appendix B) that

\[
I(\mid | Q_{\nu_i} \mid) = \begin{cases} 
I(\mid | P \mid) - \Gamma_{\nu_i} & \text{if } I(\mid | P \mid) < \infty \\
+\infty & \text{otherwise.}
\end{cases}
\]

We could proof (see the proof of lemma 4.4) that there exists a finite constant \( C \) such that

\[
|\Gamma_{\nu_i}(\mu) - \Gamma(\mu)| \leq C(1 + I(\mu \mid P)) dt(\nu_i, \mu)
\]  \( (18) \)

So that \( (17) \) implies:

\[
\lim_{N \to \infty} \frac{1}{N} \log Q^N_{\beta} (\hat{\mu}^N \in K) \leq - \inf \frac{1}{p} \inf_{K} \{(1 - C\delta)I(\mid | P \mid) - \Gamma\} + \frac{1}{q} C_q(\delta) + C\delta
\]

Finally, we proved in [2] that there exists \( \alpha < 1 \) and a finite constant \( C \) such that \( \Gamma \leq \alpha I(\mid | P \mid) + C \) so that:

\[
\liminf_{\delta \to 0} \frac{1}{p} \inf_{K} \{(1 - C\delta)I(\mid | P \mid) - \Gamma\} = \inf_{K} \{I(\mid | P \mid) - \Gamma\} = \inf_{K} H
\]

So that, letting \( \delta \downarrow 0 \), and then \( p \downarrow 1 \), we get:

\[
\lim_{N \to \infty} \frac{1}{N} \log Q^N_{\beta} (\hat{\mu}^N \in K) \leq - \inf_{K} H.
\]
5 Annealed Tightness

In this section, we prove that the law of the empirical measure under \( Q_\beta^N \) is tight, i.e. theorem 2.4:

**Theorem 5.1** For any real number \( \epsilon > 0 \), there exists a compact subset \( K_\epsilon \) of \( \mathcal{M}_1^+(W_T^A) \) such that, for any integer number \( N \),

\[
Q_\beta^N \left( \hat{\mu}_N \in K_\epsilon \right) \leq \epsilon
\]

**Proof.** To prove theorem 5.1, we shall compare the annealed law \( Q_\beta^N \) and the law of the system without interaction \( P^{\otimes N} \). To this end, first recall that, by definition of the relative entropy, for any integer \( N \), for any bounded measurable function \( f \) on \((W_T^A)^N\),

\[
\int f dQ_\beta^N \leq I(Q_\beta^N | P^{\otimes N}) + \log \int \exp \{ f \} dP^{\otimes N}
\]

Letting \( A \) be a measurable subset of \((W_T^A)^N\) and taking \( f = \log \left( 1 + P^{\otimes N}(A)^{-1} \right) I_A \), one finds that

\[
Q_\beta^N (A) \leq \frac{\log 2 + 1 + I(Q_\beta^N | P^{\otimes N})}{\log (1 + P^{\otimes N}(A)^{-1})}
\]

But the law of the empirical measure under \( P^{\otimes N} \) is exponentially tight (see lemma 3.2.7 of [4]) so that, for any real number \( \epsilon > 0 \), we can find a compact subset \( K_\epsilon \) of \( \mathcal{M}_1^+(W_T^A) \) such that

\[
P^{\otimes N} \left( \hat{\mu}_N \in K_\epsilon \right) \leq \exp \left\{ -\frac{N}{\epsilon} \right\}
\]

(19) and (20) imply that, for any real number \( \epsilon > 0 \),

\[
Q_\beta^N \left( \hat{\mu}_N \in K_\epsilon \right) \leq \frac{\log 2 + 1 + I(Q_\beta^N | P^{\otimes N})}{\log \left( 1 + \exp \left\{ \frac{N}{\epsilon} \right\} \right)}
\]

Thus, (21) implies theorem 5.1 as soon as we have proved that there exists a finite constant \( C \) such that, for any integer number \( N \),

\[
I(Q_\beta^N | P^{\otimes N}) \leq CN
\]

To compute \( I(Q_\beta^N | P^{\otimes N}) \), we recall that we proved in proposition lemma 3.6 of [2] that Girsanov Theorem implies that \( Q_\beta^N \) is absolutely continuous with respect to \( P^{\otimes N} \) and that its Radon-Nykodim derivative is given by:

\[
\frac{dQ_\beta^N}{dP^{\otimes N}} = \prod_{i=1}^N \mathcal{E} \left[ \exp \{ \beta \int_0^T G_i^{\mu_N} dB_i - \frac{\beta^2}{2} \int_0^T (G_i^{\mu_N})^2 dt \} \right]
\]

where \( G_i^{\mu_N} = \frac{1}{\sqrt{N}} \sum_{i=1}^N J_i x_i^i \) and \( \mathcal{E} \) denotes the expectation on the i.i.d. \( N(0, 1) \) random variables \( J_i \). (Remark here that \( G_i^{\mu_N} \) depends on the \( (x_i^i) \), even if we do not underline it in the notations.)
Thus, by definition of the relative entropy and of $Q_\beta^N$, we have:

\[
I(Q_\beta^N \mid P^\otimes N) = \int \log \frac{dQ_\beta^N}{dP^\otimes N} dQ_\beta^N
\]

\[
= \int \sum_{i=1}^{N} \log \mathcal{E} \left[ \exp \{ \beta \int_0^T G_t dW_t - \frac{\beta^2}{2} \int_0^T (G_t)^2 dt \} \right] dQ_\beta^N
\]

Since $Q_\beta^N$ is exchangeable, we find:

\[
I(Q_\beta^N \mid P^\otimes N) = N \int \log \mathcal{E} \left[ \exp \{ \beta \int_0^T G_t dW_t - \frac{\beta^2}{2} \int_0^T (G_t)^2 dt \} \right] dQ_\beta^N
\]  \hspace{1cm} (24)

We now give another formula for the right hand side of (24). Namely:

\[
\begin{align*}
\mathcal{E} \left[ \exp \{ \beta \int_0^T G_t dW_t - \frac{\beta^2}{2} \int_0^T (G_t)^2 dt \} \right] \\
= \exp \left\{ \beta^2 \int_0^T \mathcal{E}[\Lambda_t^N G_t dW_t] dB_t - \frac{\beta^4}{2} \int_0^T \mathcal{E}[\Lambda_t^N G_t dW_t] (G_t)^2 dt \right\}
\end{align*}
\]

\[
\text{Where} \\
\Lambda_t^N = \frac{\exp \left\{ -\frac{\beta^2}{2} \int_0^T (G_u^N)^2 du \right\}}{\mathcal{E} \left[ \exp \left\{ -\frac{\beta^2}{2} \int_0^T (G_u^N)^2 du \right\} \right]}
\]

The equality (25) is due to standard gaussian computations and integration by parts formula (see lemma 5.14 in [2]).

Thus, (24) and (25) imply that:

\[
I(Q_\beta^N \mid P^\otimes N) = N \int \left\{ \beta^2 \int_0^T \mathcal{E}[\Lambda_t^N G_t dW_t] dB_t - \frac{\beta^4}{2} \int_0^T \mathcal{E}[\Lambda_t^N G_t dW_t] (G_t)^2 dt \right\} dQ_\beta^N
\]  \hspace{1cm} (26)

Moreover, Girsanov theorem implies that, under $Q_\beta^N$, $B^1$ is a semimartingale, more precisely that there exists a $Q_\beta^N$ brownian motion $W^1$ such that, for any time $t \leq T$:

\[
B_t^1 = W_t^1 + \beta^2 \int_0^t \mathcal{E}[\Lambda_s^N G_s dW_s] dB_s - \beta^4 \int_0^t \mathcal{E}[\Lambda_s^N G_s dW_s] (G_s)^2 ds
\]  \hspace{1cm} (27)

So that (26) becomes:

\[
I(Q_\beta^N \mid P^\otimes N) = \frac{1}{2} \beta^4 N \int \int \mathcal{E} \left[ \Lambda_t^N G_t dW_t \int_0^t G_s dW_s \right]^2 dt dQ_\beta^N
\]  \hspace{1cm} (28)

14
We now bound \( f(t) = \int \mathcal{E} \left[ \Lambda_i^N G_i^N \int_0^t G_s^N dB_s \right]^2 dQ_\beta^N \) through a Gronwall lemma argument. Using (27), one finds that:

\[
f(t) = \int \mathcal{E} \left[ \Lambda_i^N G_i^N \int_0^t G_s^N dB_s \right]^2 dQ_\beta^N \\
\leq 2 \int \mathcal{E} \left[ \Lambda_i^N G_i^N \int_0^t G_s^N dW_s \right]^2 dQ_\beta^N \\
+ 2\beta^4 \int \left( \int_0^t \mathcal{E} \left[ \Lambda_i^N G_i^N \right] \mathcal{E} \left[ \Lambda_s^N G_s^N \int_0^s G_s^N dB_s \right] ds \right)^2 dQ_\beta^N \tag{29}
\]

But Cauchy Schwartz inequality in the first term in the right hand side of (29) gives:

\[
\int \mathcal{E} \left[ \left( \Lambda_i^N G_i^N \right)^2 \right] dQ_\beta^N \leq \int \mathcal{E} \left[ \left( \Lambda_i^N G_i^N \right)^2 \right] \mathcal{E} \left[ \int_0^t G_s^N dW_s \right]^2 dQ_\beta^N \tag{30}
\]

Moreover, classical gaussian properties imply (see appendix A of [2] for details), that, for any \( x \in (W_T^A)^N \):

\[
\mathcal{E} \left[ \left( \Lambda_i^N G_i^N \right)^2 \right] \leq \mathcal{E} \left[ (G_i^N)^2 \right] = \frac{1}{N} \sum_{i=1}^N (x_i')^2 \leq A^2 \tag{31}
\]

So that (30) is bounded, for any \( t \leq T \):

\[
\int \mathcal{E} \left[ \left( \Lambda_i^N G_i^N \right)^2 \right] dQ_\beta^N \leq A^2 \int \left[ \int_0^t G_s^N dW_s \right]^2 dQ_\beta^N \\
= A^2 \int \mathcal{E} \left[ \int_0^t (G_s^N)^2 ds \right] dQ_\beta^N \leq A^4 T \tag{32}
\]

Similarly, we can bound the second term in the right hand side of (29) and finally get:

\[
f(t) \leq 2A^4 T + 2\beta^4 A^4 T \int_0^t f(s) ds
\]

Since this inequality holds for any \( t \leq T \), Gronwall lemma gives:

\[
\sup_{t \leq T} f(t) = \sup_{t \leq T} \int \mathcal{E} \left[ \Lambda_i^N G_i^N \int_0^t G_s^N dB_s \right]^2 dQ_\beta^N \leq 2A^4 T \exp\{2\beta^4 A^4 T^2\}
\]

Thus, (28) implies that:

\[
I(Q_\beta^N \| P^{\otimes N}) \leq \left( \beta^4 A^4 T^2 \exp\{2\beta^4 A^4 T^2\} \right) N \tag{33}
\]

Which is the bound (22) we needed to get theorem 5.1.
References


