

# Moderate Deviations for Mean Field Particle Models

Pierre Del Moral, Shulan Hu, Liming Wu

► **To cite this version:**

Pierre Del Moral, Shulan Hu, Liming Wu. Moderate Deviations for Mean Field Particle Models. [Research Report] 2012, pp.37. <hal-00687827>

**HAL Id: hal-00687827**

**<https://hal.inria.fr/hal-00687827>**

Submitted on 15 Apr 2012

**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.

# Moderate Deviations for Mean Field Particle Models

Pierre Del Moral\*, Shulan Hu†, Liming Wu‡

April 15, 2012

## Abstract

This article is concerned with moderate deviation principles of a general class of mean field type interacting particle models. We discuss functional moderate deviations of the occupation measures for both the strong  $\tau$ -topology on the space of finite and bounded measures as well as for the corresponding stochastic processes on some class of functions equipped with the uniform topology. Our approach is based on an original semigroup analysis combined with stochastic perturbation techniques and projective limit large deviation methods.

*Keywords* : Moderate deviations, interacting particle systems, exponential inequalities, functional central limit theorems, convergence of empirical processes, large deviations for projective limits.

*MSC 2000* : Primary: 60F10 ; Secondary: 60K35.

## 1 Introduction

### 1.1 Mean Field Particle Models

Let  $(E_n)_{n \geq 0}$  be a sequence of measurable spaces equipped with some  $\sigma$ -fields  $(\mathcal{E}_n)_{n \geq 0}$ , and we let  $\mathcal{P}(E_n)$  be the set of all probability measures over the set  $E_n$ , with  $n \geq 0$ . We consider a collection of transformations  $\Phi_n : \mathcal{P}(E_{n-1}) \rightarrow \mathcal{P}(E_n)$  and we denote by  $(\eta_n)_{n \geq 0}$  a sequence of probability measures on  $E_n$  that satisfies a nonlinear equation of the following form

$$\eta_{n+1} = \Phi_{n+1}(\eta_n). \quad (1.1)$$

The mean field particle interpretations of these measure valued models relies on the fact that the one step mappings can be rewritten in the following form

$$\Phi_n(\eta_{n-1}) = \eta_{n-1} K_{n, \eta_{n-1}} \quad (1.2)$$

for some collection of Markov kernels  $K_{n, \mu_{n-1}}$  indexed by the time parameter  $n$  and the set of probability measures  $\mu_n$  on the space  $E_{n-1}$ . These models provide a natural interpretation

---

\*Centre INRIA Bordeaux Sud-Ouest Institut de Mathématiques, Université Bordeaux I, 33405 Talence cedex, France. pierre.del-moral@inria.fr.

†Corresponding author: Department of Statistics, Zhongnan University of Economics and Law, Wuhan 430073, China. hu\_shulan@yahoo.com.

‡Laboratoire de Mathématiques Appliquées, CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubiere, France. li-ming.wu@math.univ-bpclermont.fr. Institute of Applied Mathematics, Chinese Academy of Sciences, 100190 Beijing, China. wuliming@amt.ac.cn.

of the distribution laws  $\eta_n$  as the laws of a non linear Markov chain whose elementary transitions depend on the current distribution. In the further development of the article, we always assume that the mappings

$$(x_n^i)_{1 \leq i \leq N} \in E_n^N \mapsto K_{n+1, \frac{1}{N} \sum_{j=1}^N \delta_{x_n^j}} (x_n^i, A_{n+1})$$

are  $\mathcal{E}_n^{\otimes N}$ -measurable, for any  $n \geq 0$ ,  $1 \leq i \leq N$ , and any measurable subset  $A_{n+1} \subset E_{n+1}$ . In this situation, the mean field particle interpretation of this nonlinear measure valued model is an  $E_n^N$ -valued Markov chain  $\xi_n^{(N)} = (\xi_n^{(N,i)})_{1 \leq i \leq N}$ , with elementary transitions defined as

$$\mathbb{P} \left( \xi_{n+1}^{(N)} \in dx \mid \mathcal{A}_n^{(N)} \right) = \prod_{i=1}^N K_{n+1, \eta_n^N} (\xi_n^{(N,i)}, dx^i) \quad \text{with} \quad \eta_n^N := \frac{1}{N} \sum_{j=1}^N \delta_{\xi_n^{(N,j)}}. \quad (1.3)$$

In the above displayed formula,  $\mathcal{A}_n^{(N)} := \sigma \left( \xi_p^{(N)}, 0 \leq p \leq n \right)$  stands for the sigma-field generated by the random variables  $(\xi_p^{(N)})_{0 \leq p \leq n}$ , and  $dx = dx^1 \times \dots \times dx^N$  stands for an infinitesimal neighborhood of a point  $x = (x^1, \dots, x^N) \in E_n^N$ . The initial system  $\xi_0^{(N)}$  consists of  $N$  independent and identically distributed random variables with common law  $\eta_0$ . To simplify the presentation, when there is no possible confusion we suppress the parameter  $N$ , so that we write  $\xi_n$  and  $\xi_n^i$  instead of  $\xi_n^{(N)}$  and  $\xi_n^{(N,i)}$ . For a thorough description of these discrete generation and non linear McKean type models, we refer the reader to [5].

A typical example we have in mind is the Feynman-Kac model associated with  $(0, 1]$ -valued potential functions  $G_n$  and Markov transitions  $M_{n+1}$  from  $E_n$  into  $E_{n+1}$  given by

$$\Phi_{n+1}(\eta_n)(dy) = (\Psi_{G_n}(\eta_n) M_{n+1})(dy) := \int \Psi_{G_n}(\eta_n)(dx) M_{n+1}(x, dy) \quad (1.4)$$

where  $\Psi_G(\eta)(dx) := \frac{G(x)}{\eta(G)} \eta(dx)$  (with  $\eta(G) := \int G d\eta(x)$ ).

In this situation, the flow of measures  $\eta_n$  is given for any bounded measurable function  $f$  on  $E_n$  by the following functional formula

$$\eta_n(f_n) = \int_{E_n} f_n(x) \eta_n(dx) \propto \mathbb{E} \left( f_n(X_n) \prod_{0 \leq p < n} G_p(X_p) \right)$$

where  $X_n$  stands for a Markov chain with initial distribution  $\eta_0$  and Markov transitions  $M_n$ .

Recall that  $\Psi_{G_n}(\eta_n)$  can be expressed as a non-linear Markov transport equation

$$\Psi_{G_n}(\eta_n) = \eta_n S_{\eta_n, G_n} \quad (1.5)$$

with the Markov transitions

$$S_{\eta_n, G_n}(x, dy) = G_n(x) \delta_x(dy) + (1 - G_n(x)) \Psi_{G_n}(\eta_n)(dy)$$

we find that

$$K_{n+1, \eta_n} = S_{\eta_n, G_n} M_{n+1}$$

These measure valued equations arise in a variety of applications areas, including in physics, biology and in advanced stochastic engineering sciences. For instance, in signal processing, the conditional distributions of the paths of Markov signal given a series of noisy observations satisfy a two-step prediction-updating equation of the form (1.1). In this context, the state space  $E_n$  depends on the time parameter and it consists of all signal path sequences all length  $n$ . In this situation, it is worth mentioning that the corresponding mean field particle model in path space represents the evolution of a genealogical tree model associated with a genetic type algorithm.

In the context of sequential bayesian inference, the distributions  $\eta_n$  could also be the posterior distributions of an unknown parameter given the data collected up to time  $n$ . These equations also arise in physics and in molecular chemistry. In this situation,  $\eta_n$  is often interpreted as the law of a particle evolving in an absorbing medium related to some potential energy function. These non linear models are also used in advanced stochastic engineering sciences, and more particularly in stochastic optimization as well as in rare event simulation. In these situations,  $\eta_n$  is often given by a Boltzmann-Gibbs measure associated with some decreasing temperature parameter or some decreasing sequence of critical rare event levels. In the late case, the state spaces  $E_n$  represent the set of excursions of the reference Markov chain between two level sets.

During the last two decades, the mean field particle interpretations of these discrete generation measure valued equations are increasingly identified as a powerful stochastic simulation algorithm. They have led to spectacular results in signal processing with the corresponding particle filter technology, in stochastic engineering with interacting type Metropolis and Gibbs sampler methods, as well as in quantum chemistry with quantum and diffusion Monte Carlo algorithms leading to precise estimates of the top eigenvalues and the ground states of Schroedinger operators. It is clearly out of the scope of this article to review these models. For a more detailed discussion on these application areas, we again refer the reader to [5, 8, 10, 7], and the references therein.

The mathematical and numerical analysis of these mean field particle models (1.3) is one of the most active research subject in pure and applied probability, as well as in advanced stochastic engineering and computational physics. In recent years, a variety of mathematical results have been discussed in the literature, including propagation of chaos type properties,  $\mathbb{L}_p$ -mean error bounds, as well as fluctuations theorems, large deviation principles and non asymptotic concentration inequalities. The moderate deviation properties can be thought as an intermediate asymptotic estimation between the central limit theorem and the large deviations principles. Most of the existing literature on moderate deviation principles is concerned with independent and identically distributed random sequences or Markov chain processes; see for instance the series of works by M.A. Arcones [1, 2], A. de Acosta [3], A. de Acosta and X. Chen [4], H. Djellout and A. Guillin [13], F.Q. Gao [11, 12], M. Ledoux [14], L.M. Wu [19, 18].

Surprisingly very few articles discuss moderate deviations for mean field interacting particle models. The first pioneering article discussing moderate deviations for interacting processes seems to be the work by R. Douc, A. Guillin and J. Najim [9]. In this article, the authors prove a moderate deviation for the empirical mean value of a particle filter model associated with some bounded and unbounded fixed sequence of test functions. In our framework, we also mention that the particle filter stochastic model discussed in [9] is associated with a class of McKean type transitions of the form  $K_{n,\eta}(x, dy) = \Phi_n(\eta)(dy)$ .

The main simplification due to this choice of transition comes from the fact that the corresponding mean field particle model reduces to a collection of conditionally independent and identically distributed random variables.

The rather weak regularity properties used in this work follow a recent article of the first author with E. Rio [7]. The mathematical framework developed in this recent work applies to a general class of mean field particle models, including Feynman-Kac integration models, McKean Vlasov diffusion type models, as well as McKean collision type models of gases. We generalized the classical Hoeffding, Bernstein and Bennett inequalities for independent random sequences to interacting particle systems but we left open the question of moderate deviation principles. In the present article, we complete this study with functional moderate deviations of mean field particle models for both the  $\tau$ -topology on the space of signed and bounded measures and for the empirical random field processes associated with some collection of functions. Our analysis is based on an original semigroup analysis combined with stochastic perturbation techniques and projective limit deviation methods.

## 1.2 Outline of the paper.

This paper is organized as follows. In the next section we present the main results : the moderate deviation principles (MDP in short) in three types : (1) in finite dimension ; (2) in infinite dimension but for the  $\tau$ -topology; (3) for empirical process indexed by a class of functions; and we describe some main lines leading to them. We prove the MDP in finite dimension in section 3. We prove in Section 4 the MDP in the  $\tau$ -topology by the method of projective limit. We establish in Section 5 the MDP for empirical processes by the method of metric entropy. Some complicated and technical results needed in the proofs of the main theorems are provided in the three Appendices : Section 6, 7 and 8.

## 1.3 Some notations

We end this introduction with some more or less traditional notations used in the present article. We denote respectively by  $\mathcal{M}(E)$ ,  $\mathcal{M}_0(E)$ , and  $\mathcal{B}(E)$ , the set of all finite signed measures on some measurable space  $(E, \mathcal{E})$ , the convex subset of finite signed measures  $\nu$  with  $\nu(E) = 0$ , and the Banach space of all bounded and measurable functions  $f$  equipped with the uniform norm  $\|f\|$ . We also denote by  $\text{Osc}_1(E)$ , the convex set of  $\mathcal{E}$ -measurable functions  $f$  with oscillations  $\text{osc}(f) := \sup_{x \neq y} |f(x) - f(y)| \leq 1$ . We let  $\mu(f) = \int \mu(dx) f(x)$ , be the Lebesgue integral of a function  $f \in \mathcal{B}(E)$ , with respect to a measure  $\mu \in \mathcal{M}(E)$ . We recall that a bounded integral operator  $M$  from a measurable space  $(E, \mathcal{E})$  into an auxiliary measurable space  $(F, \mathcal{F})$  is an operator  $f \mapsto M(f)$  from  $\mathcal{B}(F)$  into  $\mathcal{B}(E)$  such that the functions  $x \mapsto M(f)(x) := \int_F M(x, dy) f(y)$  are  $\mathcal{E}$ -measurable and bounded, for any  $f \in \mathcal{B}(F)$ . A Markov kernel is a positive and bounded integral operator  $M$  with  $M(1) = 1$ . Given a pair of bounded integral operators  $(M_1, M_2)$ , we let  $(M_1 M_2)$  the composition operator defined by  $(M_1 M_2)(f) = M_1(M_2(f))$ . For time homogenous state spaces, we denote by  $M^m = M^{m-1} M = M M^{m-1}$  the  $m$ -th composition of a given bounded integral operator  $M$ , with  $m \geq 1$ .

A bounded integral operator  $M$  from a measurable space  $(E, \mathcal{E})$  into an auxiliary measurable space  $(F, \mathcal{F})$  also generates a dual operator  $\mu \mapsto \mu M$  from  $\mathcal{M}(E)$  into  $\mathcal{M}(F)$  defined

by  $(\mu M)(f) := \mu(M(f))$ . We let  $b(m)$  be the collection of constants given below

$$b(2m)^{2m} := \frac{(2m)!}{m!2^m}, \quad \text{and} \quad b(2m+1)^{2m+1} := \frac{(2m+1)!}{(m+1)!\sqrt{m+1/2}} 2^{-(m+1/2)}.$$

When the bounded integral operator  $M$  has a constant mass, that is  $M(1)(x) = M(1)(y)$  for any  $(x, y) \in E^2$ , the operator  $\mu \mapsto \mu M$  maps  $\mathcal{M}_0(E)$  into  $\mathcal{M}_0(F)$ . In this situation, we let  $\beta(M)$  be the Dobrushin coefficient of a bounded integral operator  $M$  defined by the following formula

$$\beta(M) := \sup \{ \text{osc}(M(f)) ; f \in \text{Osc}_1(F) \}. \quad (1.6)$$

Finally, we let  $\Phi_{p,n}$ ,  $0 \leq p \leq n$ , be the semigroup associated with the measure valued equation defined in (1.1).

$$\Phi_{p,n} = \Phi_n \circ \Phi_{n-1} \circ \dots \circ \Phi_{p+1}.$$

For  $p = n$ , we use the convention  $\Phi_{n,n} = Id$ , the identity operator.

## 2 Description of the main results and a first order fluctuation analysis

### 2.1 Regularity conditions

We let  $\Upsilon(E_1, E_2)$  be the set of mappings  $\Phi : \mu \in \mathcal{P}(E_1) \mapsto \Phi(\mu) \in \mathcal{P}(E_2)$  satisfying the first order decomposition

$$\Phi(\mu) - \Phi(\eta) = (\mu - \eta)D_\eta\Phi + \mathcal{R}^\Phi(\mu, \eta) \quad (2.1)$$

where

- (i) the first order operators  $(D_\eta\Phi)_{\eta \in \mathcal{P}(E_1)}$  is some collection of bounded integral operators from  $E_1$  into  $E_2$  such that  $\forall \eta \in \mathcal{P}(E_1)$ ,  $\forall x \in E_1$ ,  $(D_\eta\Phi)(1)(x) = 0$  and

$$\beta(\mathcal{D}\Phi) := \sup_{\eta \in \mathcal{P}(E_1)} \beta(D_\eta\Phi) < \infty; \quad (2.2)$$

- (ii) the collection of second order remainder signed measures  $(\mathcal{R}^\Phi(\mu, \eta))_{(\mu, \eta) \in \mathcal{P}(E_1^2)}$  on  $E_2$  are such that

$$|\mathcal{R}^\Phi(\mu, \eta)(f)| \leq \int |(\mu - \eta)^{\otimes 2}(g)| R_\eta^\Phi(f, dg) \quad (2.3)$$

for some collection of integral operators  $R_\eta^\Phi$  from  $\mathcal{B}(E_2)$  into the set  $\text{Osc}_1(E_1)^2$  such that

$$\sup_{\eta \in \mathcal{P}(E_1)} \int \text{osc}(g_1) \text{osc}(g_2) R_\eta^\Phi(f, d(g_1 \otimes g_2)) \leq \text{osc}(f) \delta(R^\Phi) \quad \text{with} \quad \delta(R^\Phi) < \infty. \quad (2.4)$$

We say that a collection of Markov transitions  $K_\eta$  from a measurable space  $(E_1, \mathcal{E}_1)$  into another  $(E_2, \mathcal{E}_2)$  satisfy condition (K) as soon as the following Lipschitz type inequality is met for every  $f \in \text{Osc}_1(E_2)$ :

$$(K) \quad \| [K_\mu - K_\eta](f) \| \leq \int |(\mu - \eta)(h)| T_\eta^K(f, dh). \quad (2.5)$$

In the above display,  $T_\eta^K$  stands for some collection of bounded integral operators from  $\mathcal{B}(E_2)$  into  $\mathcal{B}(E_1)$  such that

$$\sup_{\eta \in \mathcal{P}(E_1)} \int \text{osc}(h) T_\eta^K(f, dh) \leq \text{osc}(f) \delta(T^K) \quad (2.6)$$

for some finite constant  $\delta(T^K) < \infty$ . In the special case where  $K_\eta(x, dy) = \Phi(\eta)(dy)$ , for some mapping  $\Phi : \eta \in \mathcal{P}(E_1) \mapsto \Phi(\eta) \in \mathcal{P}(E_2)$ , condition (K) is a simple Lipschitz type condition on the mapping  $\Phi$ . In this situation, we denote by  $(\Phi)$  the corresponding condition; and whenever it is met, we say that the mapping  $\Phi$  satisfy condition  $(\Phi)$ .

Throughout this paper we assume

**(H1)** *The given collection of McKean transitions  $K_{n,\eta}$  satisfies the Lipschitz type condition stated in (2.5) and (2.6). We also assume that the one step mappings*

$$\Phi_n : \mu \in \mathcal{P}(E_{n-1}) \longrightarrow \Phi_n(\mu) := \mu K_{n,\mu} \in \mathcal{P}(E_n)$$

*governing the equation (1.1) are chosen so that  $\Phi_n \in \Upsilon(E_{n-1}, E_n)$ , for any  $n \geq 1$ .*

Several examples of non linear semigroups satisfying these weak regularity can be found in [7], including gaussian type mean field models, and McKean velocity models of gases. We illustrate our assumptions in the context of Feynman-Kac type models. In this situation, we have the easily checked formulae

$$\begin{aligned} [\Phi_{n+1}(\mu) - \Phi_{n+1}(\eta)](f) &= \frac{1}{\mu(G_{n,\eta})} (\mu - \eta) [G_{n,\eta} M_{n+1,\eta}(f)] \\ &= (\mu - \eta) [G_{n,\eta} M_{n+1,\eta}(f)] \\ &\quad + \frac{1}{\mu(G_{n,\eta})} [\eta - \mu](G_{n,\eta}) (\mu - \eta) [G_{n,\eta} M_{n+1,\eta}(f)] \end{aligned}$$

with the functions

$$G_{n,\eta} = G_n / \eta(G_n) \quad \text{and} \quad M_{n+1,\eta}(f) := M_{n+1}(f) - \Phi_{n+1}(\eta)(f)$$

Assuming that  $g_n = \sup_{x,y} G_n(x)/G_n(y) < \infty$ , we find the Lipschitz estimates

$$|[\Phi_{n+1}(\eta) - \Phi_{n+1}(\eta)](f)| \leq g_n |(\mu - \eta) D_\eta \Phi_{n+1}(f)| \quad (2.7)$$

as well as the first order estimation

$$|[[\Phi_{n+1}(\eta) - \Phi_{n+1}(\eta)] - (\mu - \eta) D_\eta \Phi_{n+1}](f)| \leq g_n |[\eta - \mu](G_{n,\eta})| |(\mu - \eta) [D_\eta \Phi_{n+1}(f)]|$$

with the first order functional

$$D_\eta \Phi_{n+1}(f) = G_{n,\eta} M_{n+1,\eta}(f)$$

We also mention that the corresponding one step mappings  $\Phi_n(\eta) = \eta K_{n,\eta}$  and the corresponding semigroup  $\Phi_{p,n}$  satisfy condition  $(\Phi_{p,n})$  for some collection of bounded integral operators  $T_\eta^{\Phi_{p,n}}$ .

## 2.2 Description of the main results

The best way to present moderate deviations is to start with the analysis of the fluctuations of the particle occupation measures. For mean field particle models, these central limit theorems are based on a stochastic perturbation interpretation of the local sampling errors. The random fields associated with these perturbation models are defined by below.

**Definition 2.1** We let  $(V_n^N, W_n^N)$  be the sequence of random fields defined by the pair of stochastic perturbation formulae:

$$\eta_n^N = \eta_{n-1}^N K_{n, \eta_{n-1}^N} + \frac{1}{\sqrt{N}} V_n^N = \eta_n + \frac{1}{\sqrt{N}} W_n^N \quad (2.8)$$

where  $\eta_n^N = \frac{1}{N} \sum_{j=1}^N \delta_{\xi_n^N(j)}$  is the empirical distribution of  $\xi_n^N$ . The sequence  $V_n^N$  is sometimes referred as the local sampling random field model.

Notice that the centered random fields  $V_n^N$  have conditional variance functions given by

$$\mathbb{E}(V_n^N(f_n)^2 \mid \xi_{n-1}^N) = \eta_{n-1}^N \left[ K_{n, \eta_{n-1}^N} \left( (f_n - K_{n, \eta_{n-1}^N}(f_n))^2 \right) \right]. \quad (2.9)$$

To analyze the propagation properties of the sampling errors, up to a second order remainder measure, by assumption that  $\Phi_n \in \Upsilon(E_{n-1}, E_n)$ , we have the first order decomposition

$$\Phi_n(\eta) - \Phi_n(\mu) \simeq (\eta - \mu) D_\mu \Phi_n \quad (2.10)$$

with a first order integral operator  $D_\mu \Phi_n$  from  $\mathcal{B}(E_n)$  into  $\mathcal{B}(E_{n-1})$ . The precise description of these regularity properties is provided in section 2.1.

**Definition 2.2** We let  $(\mathcal{D}_{p,n})_{0 \leq p \leq n}$  be the semigroup  $\mathcal{D}_{p,n} = \mathcal{D}_{p+1} \mathcal{D}_{p+1,n}$ , associated with the integral operator  $\mathcal{D}_n = D_{\eta_{n-1}} \Phi_n$ . We use the convention  $\mathcal{D}_{n,n} = Id$ , for  $p = n$ .

Using the decomposition

$$\begin{aligned} W_n^N &= V_n^N + \sqrt{N} [\Phi_n(\eta_{n-1}^N) - \Phi_n(\eta_{n-1})] \\ &\simeq V_n^N + W_{n-1}^N D_{\eta_{n-1}} \Phi_n \implies W_n^N \simeq \sum_{p=0}^n V_p^N \mathcal{D}_{p,n} \end{aligned} \quad (2.11)$$

we proved in [7] that the sequence of random fields  $(V_n^N)_{n \geq 0}$  converges in law, as  $N$  tends to infinity, to the sequence of  $n$  independent, Gaussian and centered random fields  $(V_n)_{n \geq 0}$  with a covariance function with, for any  $f, g \in \mathcal{B}(E_n)$ , the space of the bounded and measurable real functions on  $E_n$  and  $n \geq 0$ ,

$$\mathbb{E}(V_n(f) V_n(g)) = \eta_{n-1} K_{n, \eta_{n-1}} ([f - K_{n, \eta_{n-1}}(f)] [g - K_{n, \eta_{n-1}}(g)]). \quad (2.12)$$

In addition,  $W_n^N$  converges in law, as the number of particles  $N$  tends to infinity, to a Gaussian and centered random fields

$$W_n = \sum_{p=0}^n V_p \mathcal{D}_{p,n}. \quad (2.13)$$



Concentration inequalities associated with these fluctuations theorems are developed in some details in a recent article of the first author with E. Rio [7]. In this article, we analyze asymptotic expansions for probabilities of moderate deviations. To describe with some precision our main results, we recall the definition of a large deviation principle (*abbreviate LDP*) for random variables.

**Definition 2.3** *Let  $(\alpha(N))_{N \geq 1}$  be a sequence of positive numbers such that  $\lim_{N \rightarrow \infty} \alpha(N) = \infty$ . A sequence of random variables  $\mathcal{X}^N$  with values in a topological state space  $(S, \mathcal{S})$  satisfies an LDP with speed  $\alpha(N)$  and with good rate function  $I : x \in S \mapsto I(x) \in [0, \infty]$  if the pair of conditions below are satisfied:*

- *For every finite constant  $a < \infty$ , the level sets  $\{x \in S : I(x) \leq a\}$  are compact sets.*
- *For each  $A \in \mathcal{S}$*

$$-I\left(\overset{\circ}{A}\right) \leq \liminf_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{P}(\mathcal{X}^N \in A) \leq \limsup_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{P}(\mathcal{X}^N \in A) \leq -I(\overline{A})$$

where, for a subset  $B \subset S$ , we let  $I(B) := \inf_{x \in B} I(x)$ .

A sequence of random variables  $\mathcal{Y}^N$  is said to satisfy a moderate deviation principle (*abbreviate MDP*) with good rate function  $I$  and speed  $\alpha(N)$  if sequence of random variables  $\mathcal{X}^N := \mathcal{Y}^N / \sqrt{\alpha(N)}$  satisfies an LDP with speed  $\alpha(N)$  and with good rate function  $I$ .

The first steps in the MDP analysis of the pair of random field sequences  $(V_n^N, W_n^N)$  rely on the following pair of asymptotic Laplace expansions.

**Theorem 2.4** *For any nondecreasing function  $\alpha(N)$  such that  $\lim_{N \rightarrow \infty} \frac{\alpha(N)}{N} = 0$ , any  $n \geq 0$  and any collection of functions  $f_n \in \mathcal{B}(E_n)$ , with  $n \geq 0$ , we have*

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( \exp \left\{ \sqrt{\alpha(N)} \sum_{p=0}^n V_p^N(f_p) \right\} \right) = \frac{1}{2} \sum_{p=0}^n \mathbb{E} \left( V_p(f_p)^2 \right) \quad (2.14)$$

and

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( \exp \left\{ \sqrt{\alpha(N)} W_n^N(f_n) \right\} \right) = \frac{1}{2} \mathbb{E} \left( W_n(f_n)^2 \right). \quad (2.15)$$

The detailed proof of the above theorem is provided in section 3.2 and section 3.3. We already mention that the second expansion (2.15) is a more or less direct consequence of the first one (2.14) combined with the first order decomposition (2.11).

Let us examine some direct consequences of the above theorem. For any finite subset  $\mathcal{F}_n = \{f_n^1, \dots, f_n^{d_n}\} \subset \mathcal{B}(E_n)^{d_n}$ , with  $d_n \geq 1$ , we consider the projection mapping defined by

$$\pi_{\mathcal{F}_n} : \mu \in \mathcal{M}(E_n) \mapsto \pi_{\mathcal{F}_n}(\mu) = (\mu(f))_{f \in \mathcal{F}_n} \in \mathbb{R}^{\mathcal{F}_n} \simeq \mathbb{R}^{d_n}.$$

By a theorem of J. Gartner and R.S. Ellis, using the asymptotic Laplace expansion (2.15) we prove the following corollary.

**Corollary 2.5** *The random sequence  $\pi_{\mathcal{F}_n}(W_n^N)$  satisfy an MDP principle in  $\mathbb{R}^{d_n}$  with speed  $\alpha(N)$  and with the good rate function given for any  $v \in \mathbb{R}^{d_n}$  by the following formula*

$$I_{\mathcal{F}_n}^{W_n}(v) = \sup_{u \in \mathbb{R}^{d_n}} \left( \langle u, v \rangle - \frac{1}{2} \mathbb{E} \left( \left( \sum_{i=1}^d u^i W_n(f_n^i) \right)^2 \right) \right) \quad \text{with} \quad \langle u, v \rangle := \sum_{i=1}^{d_n} u^i v^i. \quad (2.16)$$

*If the covariance matrix  $C_{\mathcal{F}_n} := \left( \mathbb{E} \left( W_n(f_n^i) W_n(f_n^j) \right) \right)_{1 \leq i, j \leq d_n}$  is invertible, then the rate function  $I_{\mathcal{F}_n}^{W_n}$  takes the form*

$$I_{\mathcal{F}_n}^{W_n}(v) = \frac{1}{2} \langle v, C_{\mathcal{F}_n}^{-1} v \rangle.$$

In much the same way, using the asymptotic Laplace expansion (2.14) we readily prove the following corollary.

**Corollary 2.6** *The random sequences  $[\pi_{\mathcal{F}_0}(V_0^N), \dots, \pi_{\mathcal{F}_n}(V_n^N)]$  satisfy a MDP principle in  $\mathbb{R}^{d_0 + \dots + d_n}$  with speed  $\alpha(N)$ , with the good rate function given for any  $v = (v_0, \dots, v_n) \in \mathbb{R}^{d_0 + \dots + d_n}$  by the following formula*

$$I_{\mathcal{F}_{[0,n]}}^{V_{[0,n]}}(v) = \sum_{p=0}^n I_{\mathcal{F}_p}^{V_p}(v_p)$$

*with the functions  $I_{\mathcal{F}_n}^{V_n}$  on  $\mathbb{R}^{d_n}$  defined as  $I_{\mathcal{F}_n}^{W_n}$  by replacing in (2.16) the field  $W_n$  by  $V_n$ .*

In the second part of the article, we strengthen these MDP in two ways. Firstly, we derive the MDP for the random fields sequences on the set of measures equipped with the  $\tau$  topology. Our main result is the following theorem.

**Theorem 2.7** *We suppose that the state spaces  $E_n$  are Polish spaces (metric, complete and separable). In this situation, the sequence of random fields  $(V_0^N, \dots, V_n^N)$  satisfy an MDP in the product space  $\prod_{p=0}^n \mathcal{M}(E_p)$  equipped with the product  $\tau$  topology, with speed  $\alpha(N)$  and with the good rate function  $I_{[0,n]}$  given for any  $\mu = (\mu_p)_{0 \leq p \leq n} \in \prod_{p=0}^n \mathcal{M}(E_p)$  by*

$$I_{[0,n]}(\mu) = \sum_{p=0}^n I_p(\mu_p)$$

*with the good rate functions  $I_n$  on  $\mathcal{M}(E_n)$  defined for any  $\mu_n \in \mathcal{M}(E_n)$  by*

$$I_n(\mu_n) = \sup_{f \in \mathcal{B}(E_n)} \left( \mu_n(f) - \frac{1}{2} \eta_{n-1} \left( K_{n, \eta_{n-1}} [f - K_{n, \eta_{n-1}}(f)]^2 \right) \right). \quad (2.17)$$

*In addition, the sequence of random fields  $W_n^N$  satisfies an MDP in  $\mathcal{M}(E_n)$  (equipped with the  $\tau$  topology), with speed  $\alpha(N)$  and with the good rate function*

$$J_n(\nu) = \inf \left\{ \sum_{p=0}^n I_p(\mu_p) : \mu \text{ s.t. } \nu = \sum_{p=0}^n \mu_p \mathcal{D}_{p,n} \right\} = \sup_{f \in \mathcal{B}(E_n)} \left( \nu(f) - \frac{1}{2} \mathbb{E} (W_n(f)^2) \right). \quad (2.18)$$

A more explicit description of the rate functions  $I_n$  in terms of integral operators norms on Hilbert spaces can be found in section 4.1 (see also section 7, in the appendix).

Our second main result is a functional moderate deviation for stochastic processes indexed by a separable collection  $\mathcal{F}_n$  of measurable functions  $f_n : E_n \rightarrow \mathbb{R}$  such that  $\|f_n\| \leq 1$ . We let  $l_\infty(\mathcal{F}_n)$  be the space of all bounded real functions  $F_n : f \in \mathcal{F}_n \mapsto F_n(f) \in \mathbb{R}$  on  $\mathcal{F}_n$  with the sup norm  $\|F_n\|_{\mathcal{F}_n} = \sup_{f_n \in \mathcal{F}_n} |F_n(f_n)|$ . Notice that this vector space is a non separable Banach space if the set of functions  $\mathcal{F}_n$  is infinite. To measure the size of a given class  $\mathcal{F}_n$ , one considers the covering numbers  $N(\epsilon, \mathcal{F}_n, L_p(\mu))$  defined as the minimal number of  $L_p(\mu)$ -balls of radius  $\epsilon > 0$  needed to cover  $\mathcal{F}_n$ . By  $\mathcal{N}(\epsilon, \mathcal{F}_n)$ ,  $\epsilon > 0$ , and by  $\mathcal{I}(\mathcal{F}_n)$  we denote the uniform covering numbers and entropy integral given by

$$\mathcal{N}(\epsilon, \mathcal{F}_n) = \sup_{\eta \in \mathcal{P}(E_n)} \{ \mathcal{N}(\epsilon, \mathcal{F}_n, \mathbb{L}_2(\eta)) \} \quad \text{and} \quad \mathcal{I}(\mathcal{F}_n) = \int_0^2 \sqrt{\log \mathcal{N}(\epsilon, \mathcal{F}_n)} d\epsilon.$$

We further assume that

$$\text{(A1)} \quad \mathcal{N}(\epsilon, \mathcal{F}_n) < \infty \text{ for any } \epsilon > 0, \text{ and } \mathcal{I}(\mathcal{F}_n) < \infty.$$

This condition implies that the set  $\mathcal{F}_n$  is totally bounded in  $L_2(\eta)$ , for any distribution  $\eta$  on  $E_n$ . Various examples of classes of functions with finite covering and entropy integral are given in the book of Van der Vaart and Wellner [20] (see for instance p. 86, p. 135, and exercise 4 on p.150).

For any  $\delta > 0$ , we also set

$$\mathcal{F}_n(\delta) := \left\{ h = (f - g) : (f, g) \in \mathcal{F}_n \text{ s.t. } \eta_n(h^2)^{1/2} \leq \delta \right\}$$

**(A2)** There exists some separable collection  $\mathcal{F}'_n$  of measurable functions  $f_n$  on  $E_n$ , s.t.  $\|f_n\| \leq 1$ , s.t.  $\mathcal{I}(\mathcal{F}'_n) < c_0(n) \mathcal{I}(\mathcal{F}_{n+1})$ , and such that for any probability measure  $\mu$ , any  $\delta > 0$ , we have

$$\|\Phi_{n+1}(\mu) - \Phi_{n+1}(\eta_n)\|_{\mathcal{F}_{n+1}(\delta)} \leq c_2(n) \|\mu - \eta_n\|_{\mathcal{F}'_n(c_1(n)\delta)}$$

for some finite constant  $c_i(n) < \infty$ ,  $i = 0, 1, 2$ , whose values only depend on the mapping  $\Phi_{n+1}$ , and on the measure  $\eta_n$ .

We illustrate this regularity condition in the context of the Feynman-Kac models presented in (1.4). Using (2.7), we find that

$$|[\Phi_{n+1}(\eta) - \Phi_{n+1}(\eta_n)](h)| \leq g_n \left| (\mu - \eta_n) \left( \frac{G_n}{\eta_n(G_n)} (M_{n+1}(h) - \eta_{n+1}(h)) \right) \right|$$

where  $g_n = \sup_{x,y} G_n(x)/G_n(y)$  and

$$\begin{aligned} \eta_n \left( \left( \frac{G_n}{\eta_n(G_n)} (M_{n+1}(h) - \eta_{n+1}(h)) \right)^2 \right) &\leq g_n \eta_n \left( \frac{G_n}{\eta_n(G_n)} ((M_{n+1}(h) - \eta_{n+1}(h)))^2 \right) \\ &\leq g_n \eta_{n+1}(h^2) \end{aligned}$$

Using elementary manipulations, we show that **(A2)** is met with the constants  $c_1(n) = 1/(2\sqrt{g_n}) \leq 1$ ,  $c_2(n) = 2g_n^2$  and the class of functions

$$\mathcal{F}'_n = \left\{ \frac{1}{2g_n} \frac{G_n}{\eta_n(G_n)} (M_{n+1}(f) - \eta_{n+1}(f)) : f \in \mathcal{F}_{n+1} \right\}$$

Using lemma 2.3 in [6], we also prove that  $I(\mathcal{F}'_n) < c_0(n) I(\mathcal{F}_{n+1})$  for some finite constant whose values only depends on  $g_n$ .

For any finite subset  $\mathcal{G}_n \subset \mathcal{F}_n$ , we let

$$\pi_{\mathcal{F}_n, \mathcal{G}_n} : v \in l_\infty(\mathcal{F}_n) \mapsto \pi_{\mathcal{F}_n, \mathcal{G}_n}(v) = (v(g))_{g \in \mathcal{G}_n} \in l_\infty(\mathcal{G}_n) = \mathbb{R}^{\mathcal{G}_n}$$

be the restriction mapping defined by  $\pi_{\mathcal{F}_n, \mathcal{G}_n}(v)(g) = v(g)$ , for any  $g_n \in \mathcal{G}_n$ . The MDP of the stochastic processes  $W_n^N$  on  $\mathcal{L}_\infty(\mathcal{F}_n)$  are described below.

**Theorem 2.8** *Assume that the class of observables  $\mathcal{F}_n$  satisfies (A1), and condition (A2) is met. The sequence of stochastic processes  $W_n^N$  satisfy the large deviation principle in  $\mathcal{L}_\infty(\mathcal{F}_n)$  with the good rate function  $I_{\mathcal{F}_n}^{W_n}$  given below*

$$\begin{aligned} v \in \mathcal{L}_\infty(\mathcal{F}_n) \quad I_{\mathcal{F}_n}^{W_n}(v) &= \sup \left\{ I_{\mathcal{G}_n}^{W_n}(\pi_{\mathcal{F}_n, \mathcal{G}_n}(v)) : \mathcal{G}_n \subset \mathcal{F}_n, \text{ with } \mathcal{G}_n \text{ finite} \right\} \\ &= \inf \{ J_n(\nu) | \nu \in M_0(E_n), \nu(f) = v(f), \forall f \in \mathcal{F}_n \}. \end{aligned}$$

where  $J_n$  is given in (2.18).

For finite sets  $\mathcal{F}_n$ , the above theorem clearly reduces to the MDP presented in (2.16). Also observe that the  $\tau$ -topology on  $\mathcal{M}(E_n)$  is sometimes finer than the topology associated with the seminorm  $\|\mu - \eta\|_{\mathcal{F}_n}$  induced by  $\mathcal{F}_n$ . For instance, when  $E = \mathbb{R}^d$  and  $\mathcal{F} = \{1_{(-\infty, x]} ; x \in \mathbb{R}^d\}$ , the topology induced by the supremum distance

$$\|\mu - \eta\|_{\mathcal{F}} = \sup_{x \in \mathbb{R}^d} |\mu((-\infty, x]) - \eta((-\infty, x])|$$

is strictly coarser than the  $\tau$ -topology. In this situation, Theorem 2.8 is a direct consequence of Theorem 2.7. In more general situations, by [18] or a theorem of M. A. Arcones (see for instance theorem 3.2 in [1]), the MDP for stochastic processes  $W_n^N$  in  $\mathcal{L}_\infty(\mathcal{F}_n)$  is deduced from the MDP of the finite marginals  $\pi_{\mathcal{F}_n, \mathcal{G}_n}(W_n^N)$  plus the following exponential asymptotic equicontinuity condition:

$$\forall y > 0, \quad \lim_{\delta \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{P} \left( \frac{1}{\sqrt{\alpha(N)}} \|W_n^N\|_{\mathcal{F}_n(\delta)} > y \right) = -\infty$$

with the collection of functions

$$\mathcal{F}_n(\delta) := \{h_n : h_n = (f_n - g_n) \text{ with } (f_n, g_n) \in \mathcal{F}_n^2 \text{ and } \eta_n(h_n^2) \leq \delta\}.$$

### 3 Asymptotic Laplace expansions

#### 3.1 Some preliminary results

**Lemma 3.1** *For any  $0 \leq p \leq n$ , we have  $\Phi_{p,n} \in \Upsilon(E_p, E_n)$  with the first order decomposition type formula*

$$\Phi_{p,n}(\eta) - \Phi_{p,n}(\mu) = [\eta - \mu]D_\mu \Phi_{p,n} + \mathcal{R}^{\Phi_{p,n}}(\eta, \mu) \quad (3.1)$$

for some collection of bounded integral kernels  $D_\mu \Phi_{p,n}$  from  $E_p$  into  $E_n$  and some second order remainder signed measures  $\mathcal{R}^{\Phi_{p,n}}(\eta, \mu)$ . In addition, for any  $N \geq 1$ , we have the first order decomposition

$$W_n^N = \sum_{p=0}^n V_p^N \mathcal{D}_{p,n} + \frac{1}{\sqrt{N}} \mathcal{R}_n^N \quad \text{with} \quad \mathcal{R}_n^N := N \sum_{p=0}^{n-1} R_{p+1}^{\Phi_{p+1}}(\eta_p^N, \eta_p) D_{p+1,n} \quad (3.2)$$

and the semigroup  $(\mathcal{D}_{p,n})_{0 \leq p \leq n}$  introduced in (2.11).

**Lemma 3.2** *For every  $f \in \text{Osc}_1(E_n)$ ,  $N \geq 1$  and any  $n \geq 0$  and  $m \geq 1$ , we have the  $\mathbb{L}_m$  estimates:*

$$\mathbb{E} \left( |V_n^N(f_n)|^m \middle| \xi_{n-1}^{(N)} \right)^{\frac{1}{m}} \leq b(m) \quad \text{and} \quad \sqrt{N} \mathbb{E} \left( |[\eta_n^N - \eta_n](f_n)|^m \right)^{\frac{1}{m}} \leq b(m) \sum_{p=0}^n \delta(T^{\Phi_{p,n}}) \quad (3.3)$$

as well as the bias estimate

$$N \left| \mathbb{E}(\eta_n^N(f_n)) - \eta_n(f_n) \right| \leq \sum_{p=0}^n \delta(R^{\Phi_{p,n}}). \quad (3.4)$$

A detailed proof of (3.1) can be found in [7]. Formula (3.2) is a direct consequence of the following inductive decomposition

$$W_n^N = V_n^N + W_{n-1}^N \mathcal{D}_n + \sqrt{N} R^{\Phi_n}(\eta_{n-1}^N, \eta_{n-1}).$$

The proof of lemma 3.2 is postponed to section 6.1 in the appendix.

#### 3.2 Second order remainder measures

This section is mainly concerned with non asymptotic Laplace estimates of the second order remainder measures introduced in lemma 3.2, namely

$$\mathcal{R}_n^N := \sqrt{N} \left[ W_n^N - \sum_{p=0}^n V_p^N \mathcal{D}_{p,n} \right].$$

**Proposition 3.3** *For every  $f \in \text{Osc}_1(E_n)$ ,  $N \geq 1$ ,  $n \geq 0$ , we have the Laplace estimates :*

$$\forall t \in [0, 1/(2r(n))] \quad \mathbb{E} \left( \exp \left( t \sqrt{N} |\mathcal{R}_n^N(f_n)| \right) \right) \leq \frac{1}{\sqrt{1 - 2r(n)t}} \quad (3.5)$$

with some finite constant  $r(n) \leq \sum_{p=0}^{n-1} \beta(\mathcal{D}_{p+1,n}) \left( \sum_{q=0}^p \delta(T^{\Phi_{q,p}}) \right)^2 \delta(R^{\Phi_{p+1}})$ .

**Proof:**

By (3.2), we have that

$$|\mathcal{R}_n^N(f_n)| \leq \sum_{p=0}^{n-1} \int \left| (V_p^N)^{\otimes 2}(g) \right| R_{\eta_p}^{\Phi_{p+1}}(f, dg).$$

Combining (3.3) with the generalized Minkowski inequality this implies that

$$\left( \mathbb{E} |\mathcal{R}_n^N(f_n)|^m \right)^{1/m} \leq b(2m)^2 r(n).$$

We end the proof of the proposition recalling that for Gaussian centered random variable with  $\mathbb{E}(X^2) = 1$  we have that  $b(2m)^{2m} = \mathbb{E}(X^{2m})$  and for any  $t \in [0, 1/2[$

$$\mathbb{E}(\exp \{tX^2\}) = \sum_{m \geq 0} \frac{t^m}{m!} b(2m)^{2m} = 1/\sqrt{1-2t}.$$

■

**Corollary 3.4** *For every  $f \in \text{Osc}_1(E_n)$ ,  $N \geq 1$ ,  $n \geq 0$ , and for every  $\epsilon > 0$ , we have*

$$\mathbb{P} \left( |\mathcal{R}_n^N(f_n)| \geq \epsilon + \frac{r(n)}{\sqrt{N}} \right) \leq 2e^{-\frac{\epsilon\sqrt{N}}{2r(n)} \{1 - \delta_n(\epsilon, N)\}} \quad \text{where} \quad \delta_n(\epsilon, N) = \frac{r(n)}{\epsilon\sqrt{N}} \log \left( 1 + \frac{\epsilon\sqrt{N}}{r(n)} \right).$$

*In particular, for any nondecreasing function  $\alpha(N)$  such that  $\lim_{N \rightarrow \infty} \frac{\alpha(N)}{N} = 0$ , we have*

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{P} \left( |\mathcal{R}_n^N(f_n)| \geq \epsilon \sqrt{\alpha(N)} \right) = -\infty. \quad (3.6)$$

*In other words, the random fields  $\frac{1}{\sqrt{\alpha(N)}} W_n^N$  and  $\frac{1}{\sqrt{\alpha(N)}} \sum_{p=0}^n V_p^N \mathcal{D}_{p,n}$  are  $\alpha(N)$ -exponentially equivalent.*

**Proof:**

Using the fact that

$$\log \mathbb{E} \left( e^{t[\mathcal{R}_n^N(f_n) - r(n)]} \right) \leq -r(n)t - \frac{1}{2} \log(1 - 2r(n)t),$$

we readily find that

$$\mathbb{P} \left( \mathcal{R}_n^N(f_n) \geq \epsilon + r(n) \right) \leq \exp \left( - \sup_{t \leq 1/2} \left\{ \frac{\epsilon}{r(n)} t + t + \frac{1}{2} \log(1 - 2t) \right\} \right).$$

Choosing  $t = \frac{1}{2} \left( 1 - \frac{1}{1+\epsilon} \right)$ , we find that

$$\mathbb{P} \left( \mathcal{R}_n^N(f_n) \geq \epsilon + r(n) \right) \leq \exp \left( - \frac{\epsilon}{2r(n)} \left\{ 1 - \frac{r(n)}{\epsilon} \log \left( 1 + \frac{\epsilon}{r(n)} \right) \right\} \right)$$

which ends the proof of the corollary. ■

We end this section with a technical transfer lemma of Laplace asymptotic expansions for arbitrary stochastic processes. The proof is elementary, so omitted.

**Lemma 3.5** *Let  $(X_N), (Y_N)$  two sequences of random variables such that for any  $\lambda \geq 0$ ,*

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\lambda \alpha(N) X_N} \right) = \Lambda(\lambda) \quad \text{and} \quad \lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\lambda \alpha(N) |X_N - Y_N|} \right) = 0$$

*for some sequence  $\alpha(N)$  increasing to infinite and some finite logarithmic moment generating function  $\Lambda(\lambda)$ . Then for all  $\lambda \geq 0$ , we have*

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} (e^{\lambda \alpha(N) Y_N}) = \Lambda(\lambda).$$

### 3.3 Asymptotic Laplace transform estimates

This section is mainly concerned with the proof of Theorem 2.4. The fluctuation properties of the first order random field sequence  $\sum_{p=0}^n V_p^N \mathcal{D}_{p,n}$  is encoded in the pair of martingale sequences defined below.

**Definition 3.6** *We associate with collection of functions  $f = (f_n)_{n \geq 0} \in \prod_{n \geq 0} \mathcal{B}(E_n)$ , the pair of  $\sigma(\xi_0^{(N)}, \dots, \xi_n^{(N)})$ -martingale sequences given below*

$$M_n^{(N)}(f) = \sum_{p=0}^n V_p^N(f_p) \quad \text{and} \quad E_n^{(N)}(f) := \frac{1}{\mathcal{Z}_n^{(N)}(f)} \exp \left\{ \sqrt{\alpha(N)} M_n^{(N)}(f) \right\}$$

*with the stochastic product*

$$\mathcal{Z}_n^{(N)}(f) := \prod_{p=1}^n \mathbb{E} \left( \exp \left\{ \sqrt{\alpha(N)} V_p^N(f_p) \right\} \mid \xi_{p-1}^{(N)} \right).$$

For every  $N \geq 1$ , we notice that the angle bracket of  $M_n^{(N)}(f)$  is given by

$$\langle M^{(N)}(f) \rangle_n = \sum_{p=0}^n \Delta_p \langle M^{(N)}(f) \rangle$$

with the random increments

$$\Delta_n \langle M^{(N)}(f) \rangle := \eta_{n-1}^N \left( K_{n, \eta_{n-1}^N} \left[ \left( f_n - K_{n, \eta_{n-1}^N}(f_n) \right)^2 \right] \right).$$

We know that the sequence of martingales  $M_n^{(N)}(f)$  converges in law, as  $N$  tends to infinity, to the Gaussian martingale

$$M_n(f) = \sum_{p=0}^n V_p(f_p) \quad \text{with} \quad \langle M(f) \rangle_n = \sum_{p=1}^n \eta_{p-1} \left( K_{p, \eta_{p-1}} \left[ \left( f_p - K_{p, \eta_{p-1}}(f_p) \right)^2 \right] \right).$$

The main object of this subsection is to prove that

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N) \left( \frac{1}{\sqrt{\alpha(N)}} M_n^{(N)}(f) \right)} \right) = \frac{1}{2} \langle M(f) \rangle_n. \quad (3.7)$$

Notice that the above asymptotic Laplace expansion is equivalent to (2.14).

The next technical lemma is pivotal.

**Lemma 3.7** *There exist a pair of functions  $(\tau_{j,n}^{(N)}(f))_{j=1,2}$  that converge to 0 as  $N$  tends to  $\infty$ , such that*

$$e^{\sqrt{\alpha(N)}M_n^{(N)}(f) - \frac{\alpha(N)}{2} \langle M^{(N)}(f) \rangle_n} \leq E_n^{(N)}(f) e^{\tau_{2,n}^{(N)}(f) \frac{\alpha(N)}{2} \langle M^{(N)}(f) \rangle_n}$$

and

$$E_n^{(N)}(f) e^{-\tau_{1,n}^{(N)}(f) \frac{\alpha(N)}{2} \langle M^{(N)}(f) \rangle_n} \leq e^{\sqrt{\alpha(N)}M_n^{(N)}(f) - \frac{\alpha(N)}{2} \langle M^{(N)}(f) \rangle_n}.$$

The proof of lemma 3.7 is rather technical, thus we postpone it to section 6.3 in the appendix.

**Proposition 3.8**

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N) \left( \frac{1}{\sqrt{\alpha(N)}} M_n^{(N)}(f) - \frac{1}{2} \langle M^{(N)}(f) \rangle_n \right)} \right) = 0$$

and

$$\mathbb{E} \left( e^{t\sqrt{N} |\langle M^{(N)}(f) \rangle_n - \langle M(f) \rangle_n|} \right) \leq (1 + t\bar{c}_n) e^{(\bar{c}_n t)^2 / 2}. \quad (3.8)$$

In the above display,  $\bar{c}_n$  stands for some finite constant  $\bar{c}_n := \sum_{p=0}^n c(p)$  with

$$c(p) := 2 \{1 + \delta(T^{\Phi_p}) + \delta(T^{K_p})\} \sum_{0 \leq q < p} \delta(T^{\Phi_{q,p-1}}).$$

Before getting into the details of the proof of the above proposition, it is convenient to make a couple of comments. Firstly, replacing in (3.8) the parameter  $t$  by  $\frac{\alpha(N)}{\sqrt{N}}t$  we find that

$$\mathbb{E} \left( e^{t\alpha(N) |\langle M^{(N)}(f) \rangle_n - \langle M(f) \rangle_n|} \right) \leq \left( 1 + \frac{t\alpha(N)}{\sqrt{N}} \bar{c}_n \right) \exp \left\{ \frac{t^2 \alpha(N)^2}{2N} \bar{c}_n^2 \right\}.$$

from which we conclude that

$$\forall t \geq 0, \quad \limsup_{N \rightarrow \infty} \frac{1}{\alpha(N)} \mathbb{E} \left( e^{\alpha(N) t |\langle M^{(N)}(f) \rangle_n - \langle M(f) \rangle_n|} \right) = 0.$$

Also observe that the stochastic processes

$$\begin{aligned} A_n^N(f) &= \frac{1}{\sqrt{\alpha(N)}} M_n^{(N)}(f) - \frac{1}{2} \langle M^{(N)}(f) \rangle_n \\ B_n^N(f) &= \frac{1}{\sqrt{\alpha(N)}} M_n^{(N)}(f) - \frac{1}{2} \langle M(f) \rangle_n \end{aligned}$$

on the set of sequence  $f = (f_p)_{0 \leq p \leq n} \in \prod_{p=0}^n \mathcal{B}(E_p)$ , have the following scaling properties

$$|A_n^N(f) - \epsilon^{-1} A_n^N(\epsilon f)| = \frac{1}{2} \langle M^{(N)}(f) \rangle_n (1 - \epsilon) \leq \frac{1}{2} (1 - \epsilon) \sum_{p=0}^n \text{osc}(f_p)^2$$

and

$$|B_n^N(f) - \epsilon^{-1} B_n^N(\epsilon f)| = \frac{1}{2} \langle M(f) \rangle_n (1 - \epsilon)$$



for any  $\epsilon \in [0, 1]$ . In the above display,  $\epsilon f$  stands for the sequence of functions  $(\epsilon f_p)_{0 \leq p \leq n}$ . Therefore the asymptotic Laplace expansion (3.7) is a direct consequence of the transfer lemma 3.5.

Now, we come to

**Proof of proposition 3.8.**

Since we have  $\langle M^{(N)}(f) \rangle_n \leq \sigma_n^2(f) := \sum_{p=0}^n \text{osc}(f_p)^2$ , using lemma 3.7 we readily prove that

$$-\tau_{1,n}^{(N)}(f) \frac{1}{2} \sigma_n^2(f) \leq \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N) \left( \frac{1}{\sqrt{\alpha(N)}} M_n^{(N)}(f) - \frac{1}{2} \langle M^{(N)}(f) \rangle_n \right)} \right) \leq \tau_{2,n}^{(N)}(f) \frac{1}{2} \sigma_n^2(f).$$

This ends the proof of the first assertion. Now, we come to the proof of (3.8). For every  $n \geq 1$ ,  $\eta \in \mathcal{P}(E_{n-1})$  and  $f_n \in \mathcal{B}(E_n)$  we set

$$\Sigma_n(\eta, f_n) := \eta \left( K_{n,\eta} \left[ (f_n - K_{n,\eta}(f_n))^2 \right] \right).$$

For  $n = 0$ , we set  $\Sigma_0(\eta, f_0) = \eta([f_0 - \eta(f_0)]^2)$ . Firstly, we observe that

$$\begin{aligned} \Sigma_n(\eta, f_n) - \Sigma_n(\mu, f_n) &= [\Phi_n(\eta) - \Phi_n(\mu)] (f_n^2) + \mu (K_{n,\mu}(f_n)^2) - \eta (K_{n,\eta}(f_n)^2) \\ &= [\Phi_n(\eta) - \Phi_n(\mu)] (f_n^2) + [\mu - \eta] (K_{n,\eta}(f_n)^2) \\ &\quad + \mu (K_{n,\mu}(f_n)^2 - K_{n,\mu}(f_n)^2). \end{aligned}$$

This implies that

$$\begin{aligned} &|\Sigma_n(\eta, f_n) - \Sigma_n(\mu, f_n)| \\ &\leq |[\Phi_n(\eta) - \Phi_n(\mu)] (f_n^2)| + |[\mu - \eta] (K_{n,\eta}(f_n)^2)| + 2\|K_{n,\mu}(f_n) - K_{n,\eta}(f_n)\| \end{aligned}$$

and therefore

$$\begin{aligned} &(\mathbb{E}|\Sigma_n(\eta_{n-1}^N, f_n) - \Sigma_n(\eta_{n-1}, f_n)|^m)^{\frac{1}{m}} \\ &\leq \int (\mathbb{E}|(\eta_{n-1}^N - \eta_{n-1})(g)|^m)^{\frac{1}{m}} T_{\eta_{n-1}}^{\Phi_n}(f_n^2, dg) + (\mathbb{E}|(\eta_{n-1}^N - \eta_{n-1})(K_{n,\eta_{n-1}}(f_n)^2)|^m)^{\frac{1}{m}} \\ &\quad + 2 \int \mathbb{E}(|(\eta_{n-1}^N - \eta_{n-1})(g)|^m)^{\frac{1}{m}} T_{\eta_{n-1}}^{K_n}(f_n, dg). \end{aligned}$$

Using (3.3), we have the upper bound

$$\sqrt{N} \mathbb{E} (|\Sigma_n(\eta_{n-1}^N, f_n) - \Sigma_n(\eta_{n-1}, f_n)|^m)^{\frac{1}{m}} \leq b(m) c(n).$$

One concludes that

$$\sqrt{N} \mathbb{E} \left( |\langle M^{(N)}(f) \rangle_n - \langle M(f) \rangle_n|^m \right)^{\frac{1}{m}} \leq b(m) \bar{c}_n.$$

The  $\mathbb{L}_m$ -inequalities stated above clearly imply that for any  $t > 0$

$$\begin{aligned}
& \mathbb{E} \left( \exp \left\{ t\sqrt{N} \left| \langle M^{(N)}(f) \rangle_n - \langle M(f) \rangle_n \right| \right\} \right) \\
&= \sum_{m \geq 0} \frac{t^{2m}}{(2m)!} \mathbb{E} \left( \left( \langle M^{(N)}(f) \rangle_n - \langle M(f) \rangle_n \right)^{2m} \right) \\
&\quad + \sum_{m \geq 0} \frac{t^{2m+1}}{(2m+1)!} \mathbb{E} \left( \left| \langle M^{(N)}(f) \rangle_n - \langle M(f) \rangle_n \right|^{2m+1} \right) \\
&\leq \sum_{m \geq 0} \frac{1}{m!} \left( \frac{t^2 \bar{c}_n^2}{2} \right)^m + (t\bar{c}_n) \sum_{m \geq 0} \frac{1}{m!} \left( \frac{t^2 \bar{c}_n^2}{2} \right)^m.
\end{aligned}$$

where (3.8) follows. ■

### 3.4 Proof of Theorem 2.4

*Proof of (2.14).* This is done in Subsection 3.3.

*Proof of (2.14)  $\implies$  (2.15).* Note that if a final time horizon  $n$  is fixed then we have for any function  $f_n \in \mathcal{B}(E_n)$

$$(\forall 0 \leq p \leq n \quad f_p = \mathcal{D}_{p,n}(f_n)) \implies \sum_{p=0}^n V_p^N(f_p) = \sum_{p=0}^n V_p^N \mathcal{D}_{p,n}(f_n).$$

Let  $(A_n^N, B_n^N)$  the pair of random fields defined below:

$$A_n^N = \frac{1}{\sqrt{\alpha(N)}} \sum_{p=0}^n V_p^N \mathcal{D}_{p,n} \quad \text{and} \quad B_n^N = \frac{1}{\sqrt{\alpha(N)}} W_n^N.$$

By (2.14), we have

$$\lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N) A_n^N(f_n)} \right) = A_n(f_n) := \mathbb{E} \left( \frac{1}{2} \sum_{p=0}^n V_p(\mathcal{D}_{p,n}(f_n))^2 \right)$$

and by (3.5)

$$\forall t \in [0, N/(2\alpha(N)r(n))[, \quad \mathbb{E} \left( e^{t \alpha(N) | [B_n^N - A_n^N](f_n) |} \right) \leq \left( 1 - \frac{\alpha(N)2r(n)t}{N} \right)^{-\frac{1}{2}}.$$

This yields that

$$\forall t > 0 \quad \lim_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{t \alpha(N) | [B_n^N - A_n^N](f_n) |} \right) = 0$$

where (2.15) follows by the transfert lemma 3.5.

## 4 Moderate deviations in $\tau$ -topology

We further require that the state spaces  $E_n$  are Polish spaces. The  $\tau$ -topology on  $\mathcal{M}(E_n)$  is the coarsest topology that makes the maps  $\mu \in \mathcal{M}(E_n) \mapsto \mu(A)$  continuous, for any measurable set  $A \in \mathcal{E}$ .

### 4.1 A deviation theorem for the local sampling random fields

The main object of this section is to prove the following theorem.

**Theorem 4.1** *The sequence of random fields  $V_n^N$  satisfy an MDP in  $\mathcal{M}(E_n)$  equipped with the  $\tau$  topology, with speed  $\alpha(N)$  and with the good rate function*

$$I_n(\mu) = \sup_{f \in \mathcal{B}(E_n)} \left( \mu(f) - \frac{1}{2} \eta_{n-1} \left( K_{n, \eta_{n-1}} [f - K_{n, \eta_{n-1}}(f)]^2 \right) \right). \quad (4.1)$$

In addition, for any  $n \geq 0$ , the sequence of random fields  $V_{[0, n]}^N := (V_0^N, \dots, V_n^N)$  satisfy an MDP in the product space  $\prod_{p=0}^n \mathcal{M}(E_n)$ , with speed  $\alpha(N)$  and with the good rate function

$$I_{[0, n]}(\mu_0, \dots, \mu_n) = \sum_{p=0}^n I_p(\mu_p).$$

Before entering into the proof of this theorem, we provide a more explicit representation of the rate functions  $I_n$ . Firstly, assume that the McKean transitions  $K_{n, \eta}$  are given by  $K_{n, \eta}(x, dy) = \Phi_n(\eta)(dy)$ . In this situation, we have

$$K_{n, \eta_{n-1}}(x, dy) = \eta_n(dy) \Rightarrow I_n(\mu) = \sup_{f \in \mathcal{B}(E_n)} \left( \mu(f) - \frac{1}{2} \eta_n \left( [f - \eta_n(f)]^2 \right) \right).$$

The variational formula given above coincides with the one of the rate function of the MDP associated with independent and identically distributed random sequences. In this case, we have that

$$I_n(\mu) = \left\| \frac{d\mu}{d\eta_n} \right\|_{\mathbb{L}_2(\eta_n)}^2 \quad \text{if } \mu \ll \eta_n \quad \text{with } \frac{d\mu}{d\eta_n} \in \mathbb{L}_2(\eta_n) \quad \text{and } \mu(E) = 0$$

and  $I_n(\mu) = \infty$ , otherwise. A proof of this assertion is provided in section 7, in the appendix. In more general situations, we need to work a little harder. Let  $K_{n, \eta_{n-1}}^*$  be the adjoint operator of  $K_{n, \eta_{n-1}}$  from  $\mathbb{L}_2(\eta_{n-1})$  into  $\mathbb{L}_2(\eta_n)$  given by

$$\forall (f, g) \in \mathbb{L}_2(\eta_n) \times \mathbb{L}_2(\eta_{n-1}), \quad \eta_n \left( f K_{n, \eta_{n-1}}^*(g) \right) = \eta_{n-1} \left( K_{n, \eta_{n-1}}(f) g \right).$$

We will prove in section 7.4. the following explicit expression:

$$I_n(\mu) = \frac{1}{2} \sum_{m \geq 0} \eta_n \left[ (h_\mu) \left( K_{n, \eta_{n-1}}^* K_{n, \eta_{n-1}} \right)^m (h_\mu) \right], \quad \mu \in \mathcal{M}_0(E_n), \mu \ll \eta_n, h_\mu = \frac{d\mu}{d\eta_n} \in \mathbb{L}_2(\eta_n) \quad (4.2)$$

and  $I_n(\mu) = +\infty$  otherwise.

## 4.2 Moderate deviations for projective limits

The proof of the theorem is based on a projective limit interpretation of the strong topology on the set of finite and signed measures. We begin by first introducing several definitions.

**Definition 4.2** We let  $\mathcal{U}(E_n)$  the set of finite partitions  $U_n = (U_n^i)_{1 \leq i \leq d} \in \mathcal{E}_n^d$  of the set  $E_n$ , with  $d \geq 1$ . We let  $\sigma(U_n)$  be the  $\sigma$ -field generated by  $U_n$ . We also let

$$\pi_{U_n} : \mu \in \mathcal{M}(E) \mapsto \pi_{U_n}(\mu) \in \mathcal{M}(E_n, \sigma(U_n))$$

be the restriction of the measure  $\mu$  to the sigma-field  $\sigma(U_n)$ .

Notice that  $\mathcal{M}(E_n, \sigma(U_n))$  can be identified with  $\mathbb{R}^{U_n} \simeq \mathbb{R}^d$ . Furthermore, the  $\sigma$ -algebra and the  $\tau$ -topology induced on  $\mathcal{M}(E_n, \sigma(U_n))$  by the restriction mapping  $\pi_{U_n}$  coincide with the natural topology and the Borel sigma-field on  $\mathbb{R}^d$ .

**Definition 4.3** We say that a partition  $U'_n$  is finer than  $U_n$ , and we write  $U'_n \geq U_n$ , as soon as we have  $\sigma(U'_n) \supset \sigma(U_n)$ . We also let  $\pi_{U'_n, U_n} : \mu \in \mathcal{M}(E, \sigma(U'_n)) \mapsto \pi_{U'_n, U_n}(\mu) \in \mathcal{M}(E_n, \sigma(U_n))$  be the restriction of the measure  $\mu$  on  $\sigma(U'_n)$  to the sigma-field  $\sigma(U_n)$ . The set  $(\mathcal{M}(E_n, \sigma(U_n)), \pi_{U'_n, U_n})_{U'_n \geq U_n}$  forms a projective inverse spectrum of  $\mathcal{U}(E_n)$ . We let  $\lim_{\mathcal{U}_n} \mathcal{M}_n$  be the projective limit space of the spectrum

$$\lim_{\mathcal{U}_n} \mathcal{M}_n := \left\{ \mu \in \prod_{U_n \in \mathcal{U}_n} \mathcal{M}(E_n, \sigma(U_n)) : \forall U'_n \geq U_n \quad \pi_{U_n}(\mu) = \pi_{U'_n, U_n}(\pi_{U'_n}(\mu)) \right\}.$$

**Definition 4.4** We let  $\mathbf{M}(E_n)$  be the set of finite additive set functions from  $\mathcal{E}_n$  into  $\mathbb{R}_+$ , equipped with the  $\tau_1$ -topology of setwise convergence. More precisely, a sequence  $\mu_k \in \mathbf{M}(E_n)$   $\tau_1$ -converges to some  $\mu \in \mathbf{M}(E_n)$  as soon as  $\lim_{k \rightarrow \infty} \mu_k(A) = \mu(A)$ , for any  $A \in \mathcal{E}_n$ .

We let  $\theta : \lim_{\mathcal{U}_n} \mathcal{M}_n \rightarrow \mathbf{M}(E_n)$  be the mapping that associates a point  $\mu = (\mu^{U_n})_{U_n \in \mathcal{U}_n} \in \lim_{\mathcal{U}_n} \mathcal{M}_n$  the set function  $\theta \in \mathbf{M}(E_n)$  defined for any  $A \in \mathcal{E}_n$  by

$$\theta(\mu)(A) = \mu^{U_n}(A) \quad \text{where } U_n \in \mathcal{U}_n \text{ is such that } A \in \sigma(U_n).$$

By construction of the projective inverse spectrum and by definition of the  $\tau_1$  convergence, it is readily checked that  $\theta$  is an homeomorphism.

By Theorem 2.4, the random sequence  $V_n^N(U_n) := (V_n^N(U_n^1), \dots, V_n^N(U_n^d))$  satisfies a MDP in  $\mathbb{R}^d$ , with speed  $\alpha(N)$  and with the good rate function

$$I_{U_n}(v^1, \dots, v^d) := \sup_{u \in \mathbb{R}^d} \left( \langle u, v \rangle - \frac{1}{2} \mathbb{E} \left( \left( \sum_{i=1}^d u^i V_n(U_n^i) \right)^2 \right) \right).$$

Since we have

$$\sum_{i=1}^d u^i V_n(U_n^i) = V_n(f_u) \quad \text{with} \quad f_u := \sum_{i=1}^d u^i 1_{U_n^i}$$

we readily find that

$$\frac{1}{2} \mathbb{E} \left( \left( \sum_{i=1}^d u^i V_n(U_n^i) \right)^2 \right) = \frac{1}{2} \eta_{n-1} \left( K_{n, \eta_{n-1}} [f_u - K_{n, \eta_{n-1}}(f_u)]^2 \right)$$

from which we conclude that

$$I_{U_n}(\pi_{U_n}(\mu)) := \sup_{f \in \mathcal{B}(E_n, \sigma(U_n))} \left( \mu(f) - \frac{1}{2} \eta_{n-1} \left( K_{n, \eta_{n-1}} [f - K_{n, \eta_{n-1}}(f)]^2 \right) \right).$$

By a theorem of D. Dawson and J. Gärtner, we deduce the following

**Proposition 4.5** *The sequence of random fields  $V_n^N$  satisfy an MDP in  $\mathbf{M}(E_n)$  ( $\simeq \lim_{\mathcal{U}_n} \mathcal{M}_n$ ), with speed  $\alpha(N)$  and with the good rate function*

$$\bar{I}_n(\mu) = \sup_{U_n \in \mathcal{U}_n} I_{U_n}(\pi_{U_n}(\mu)). \quad (4.3)$$

The proof of (4.1) is now a direct consequence of the next lemma.

**Lemma 4.6** *The domain  $\text{Dom}(\bar{I}_n) = \{\mu \in \mathbf{M}(E_n) : \bar{I}_n(\mu) < \infty\}$  of the mapping  $\bar{I}_n$  is included in  $\mathcal{M}(E_n)$  and for any  $\mu \in \mathcal{M}(E_n)$ , the rate function  $\bar{I}_n(\mu)$  defined in (4.3) coincide with  $I_n$  in (4.1).*

Before getting into the proof of the lemma, it is convenient to make some remarks.

Firstly, since the relative topology on  $\mathcal{M}(E_n)$  induced by the  $\tau_1$  topology coincide with the  $\tau$  topology, one concludes that the sequence of random fields  $V_n^N$  satisfies a MDP in  $\mathcal{M}(E_n)$  with good rate function  $I_n$ .

Furthermore, since the projection operators  $\pi_{U_n}$  are  $\tau$ -continuous, by the contraction principle one concludes that the random fields sequence  $\pi_{U_n}(V_n^N)$  satisfies a MDP in  $\mathcal{M}(E_n, \sigma(U_n))$  with the good rate function

$$I_{U_n}(\nu) := \inf \{I_n(\mu) : \mu \in \mathcal{M}(E_n) \text{ s.t. } \pi_{U_n}(\mu) = \nu\}.$$

These constructions extend in a natural way to the sequence of random fields  $(V_n^N)_{n \geq 0}$ . Indeed, using (2.14), we find that the random sequences

$$(V_0^N(U_0), \dots, V_n^N(U_n)) \quad \text{with} \quad (U_0, \dots, U_n) \in (\mathcal{U}_0 \times \dots \times \mathcal{U}_n)$$

satisfy an MDP in  $(\mathbb{R}^{d_0} \times \dots \times \mathbb{R}^{d_n})$ , with speed  $\alpha(N)$  and with the good rate function

$$I_{U_0, \dots, U_n}(v_0, \dots, v_n) := \sum_{p=0}^n \sup_{u_p \in \mathbb{R}^{d_p}} \left( \langle u_p, v_p \rangle - \frac{1}{2} \mathbb{E} (V_p(f_p^{u_p})^2) \right)$$

with the sequence of functions  $f_n^{u_n} = \sum_{i=1}^d u_n^i 1_{U_n^i}$ . The proof of theorem 2.7 is now easily completed.

Now, we come to the

**Proof of lemma 4.6 :** Consider a sequence of partitions  $U_{n,d}$ , finer and finer when  $d$  increases, such that  $\sigma\left(\bigcup_{d \geq 1} U_{n,d}\right) = \mathcal{E}_n$ . To prove that  $\text{Dom}(\bar{I}_n) \subset \mathcal{M}(E_n)$ , we use the fact that

$$I_{U_{n,d}}(\pi_{U_{n,d}}(\mu)) < \infty \Rightarrow \pi_{U_{n,d}}(\mu) \ll \pi_{U_{n,d}}(\eta_n)$$

and

$$\pi_{U_{n,d}}(\eta_n) \left( \left( \frac{d\pi_{U_{n,d}}(\mu)}{d\pi_{U_{n,d}}(\eta_n)} \right)^2 \right) \leq I_{U_{n,d}}(\pi_{U_{n,d}}(\mu)) \leq \bar{I}_n(\mu) < \infty.$$

See for instance (7.5) in the appendix. Therefore  $\left\{ \frac{d\pi_{U_{n,d}}(\mu)}{d\pi_{U_{n,d}}(\eta_n)} \right\}_{d \geq 1}$  is a  $\mathbb{L}_2$ -bounded martingale w.r.t. the probability measure  $\eta_n$  and the filtration  $(\sigma(U_{n,d}))_{d \geq 1}$ . By the martingale convergence theorem, there is some  $h_\mu \in \mathbb{L}_2(\eta_n)$  such that

$$\frac{d\pi_{U_{n,d}}(\mu)}{d\pi_{U_{n,d}}(\eta_n)} \rightarrow h_\mu$$

in  $\mathbb{L}_2(\eta_n)$ , as  $d$  goes to infinity. We show now that  $h_\mu$  does not depend on the sequence  $(U_{n,d})$ . In fact if  $(U'_{n,d})_{d \geq 1}$  is another such sequence of partitions, we consider the partition  $V_{n,d}$  which is finer than  $U_{n,d}$  and  $U'_{n,d}$  such that  $V_{n,d+1}$  is finer than  $V_{n,d}$ . By the above argument, we have

$$\frac{d\pi_{U'_{n,d}}(\mu)}{d\pi_{U'_{n,d}}(\eta_n)} \rightarrow h'_\mu, \quad \frac{d\pi_{V_{n,d}}(\mu)}{d\pi_{V_{n,d}}(\eta_n)} \rightarrow \tilde{h}_\mu$$

in  $\mathbb{L}_2(\eta_n)$ , as  $d \rightarrow \infty$ . Consequently for any  $\sigma(U_{n,d})$ -measurable and bounded function  $f$  (with  $d$  fixed),

$$\eta_n(h_\mu f) = \eta_n \left( \frac{d\pi_{U_{n,d}}(\mu)}{d\pi_{U_{n,d}}(\eta_n)} f \right) = \pi_{U_{n,d}}(\mu)(f) = \pi_{V_{n,d}}(\mu)(f) = \eta_n(\tilde{h}_\mu f).$$

Thus  $h_\mu = \tilde{h}_\mu$ ,  $\eta_n - a.s.$ . By the same way  $h'_\mu = \tilde{h}_\mu$ ,  $\eta_n - a.s.$ . Hence  $h_\mu$  does not depend on  $(U_{n,d})$ .

Finally for any finite partition  $U_n$  and  $\sigma(U_n)$ -measurable function  $f$ , taking a sequence of partitions  $(U_{n,d})$  containing  $U_n$ , we get for  $d$  large enough

$$\mu(f) = \pi_{U_{n,d}}(\eta_n) \left( \frac{d\pi_{U_{n,d}}(\mu)}{d\pi_{U_{n,d}}(\eta_n)} f \right) = \eta_n(fh_\mu).$$

Consequently  $\mu$  is the measure  $h_\mu \eta_n$ .

For the last assertion, we see that

$$\begin{aligned} \bar{I}_n(\mu) &= \sup_{U_n \in \mathcal{U}_n} \sup_{f \in \mathcal{B}(E_n, \sigma(U_n))} \left( \mu(f) - \frac{1}{2} \mathbb{E}V_n(f)^2 \right) \\ &= \sup_{f \in \bigcup_{U_n \in \mathcal{U}_n} \mathcal{B}(E_n, \sigma(U_n))} \left( \mu(f) - \frac{1}{2} \mathbb{E}V_n(f)^2 \right) \\ &= \sup_{f \in \mathcal{B}(E_n)} \left( \mu(f) - \frac{1}{2} \mathbb{E}V_n(f)^2 \right) = I_n(\mu) \end{aligned}$$

by the fact that for any  $f \in \mathcal{B}(E_n)$ , there is a sequence  $f_k \in \bigcup_{U_n \in \mathcal{U}_n} \mathcal{B}(E_n, \sigma(U_n))$  which converge uniformly to  $f$  over  $E_n$ , and  $\mathbb{E}V_n(f_k)^2 \rightarrow \mathbb{E}V_n(f)^2$  by the expression of  $\mathbb{E}V_n(f)^2$ . ■

### 4.3 Some contraction properties

By the contraction principle, the moderate deviation principles presented in Theorem 2.7 can be transferred to continuous transformations of the local sampling random fields  $V_n^N$ . For instance, we have the following proposition.

**Proposition 4.7** *The random fields  $\sum_{p=0}^n V_p^N \mathcal{D}_{p,n}$  and  $W_n^N$  satisfy the MDP in  $\mathcal{M}(E_n)$  with the good rate function*

$$\begin{aligned} J_n(\nu) &= \inf \left\{ \sum_{p=0}^n I_p(\mu_p) : (\mu_p)_{0 \leq p \leq n} \in \prod_{p=0}^n \mathcal{M}(E_p) \text{ s.t. } \nu = \sum_{p=0}^n \mu_p \mathcal{D}_{p,n} \right\} \quad (4.4) \\ &= \sup_{f \in \mathcal{B}(E_n)} \left( \nu(f) - \frac{1}{2} \mathbb{E} (W_n(f)^2) \right). \end{aligned}$$

**Proof:**

The fact that  $\sum_{p=0}^n V_p^N \mathcal{D}_{p,n}$  satisfies a MDP in  $\mathcal{M}(E_n)$  with the the good rate function (4.4) is an immediate consequence of theorem 2.7. On the other hand, using (2.14) and (2.15) we prove that both random sequences

$$W_n^N(U_n) := \left( W_n^N(U_n^1), \dots, W_n^N(U_n^d) \right)$$

and

$$\sum_{p=0}^n V_p^N \mathcal{D}_{p,n}(U_n) := \left( \sum_{p=0}^n V_p^N \mathcal{D}_{p,n}(U_n^1), \dots, \sum_{p=0}^n V_p^N \mathcal{D}_{p,n}(U_n^d) \right)$$

with  $U_n = (U_n^i)_{1 \leq i \leq d} \in \mathcal{U}_n$ , satisfies a MDP in  $\mathbb{R}^d$ , with speed  $\alpha(N)$  and with the good rate function

$$\mathcal{J}_{U_n}(v^1, \dots, v^d) := \sup_{f \in \mathcal{B}(E_n, \sigma(U_n))} \left( \mu(f) - \frac{1}{2} \mathbb{E} (W_n(f)^2) \right).$$

We conclude that both random fields  $W_n^N$  and  $\sum_{p=0}^n V_p^N \mathcal{D}_{p,n}$  satisfies the same MDP in  $\mathcal{M}(E_n)$  with the good rate function

$$\mathcal{J}_n(\nu) := \sup_{U_n \in \mathcal{U}_n} \sup_{f \in \mathcal{B}(E_n, \sigma(U_n))} \left( \nu(f) - \frac{1}{2} \mathbb{E} (W_n(f)^2) \right) = J_n(\nu).$$

The the last formula comes from the uniqueness property of the rate function. This ends the proof of the proposition.  $\blacksquare$

## 5 Moderate deviations for stochastic processes

This section is mainly concerned with the proof of theorem 2.8. By a recent theorem of M. A. Arcones (see for instance theorem 3.2 in [1]), this theorem is a direct consequence of the following lemma.

**Lemma 5.1** *Under the conditions (A1) and (A2), for any  $y > 0$  we have*

$$\lim_{\delta \rightarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{P} \left( \frac{1}{\sqrt{\alpha(N)}} \|W_n^N\|_{\mathcal{F}_n(\delta)} > y \right) = -\infty$$

with the set of functions  $\mathcal{F}_n(\delta)$  given below:

$$\mathcal{F}_n(\delta) := \{h_n : h_n = (f_n - g_n) : (f_n, g_n) \in \mathcal{F}_n^2 : \eta_n(h_n^2)^{1/2} \leq \delta\}.$$

**Proof:**

The proof of this lemma is based on several key properties of empirical processes associated with conditionally independent sequences. These results are more or less well known, thus their are housed in the appendix 8.

By construction, recalling that  $0 \in \mathcal{F}_n$ , if we choose  $\delta = 2$  then we have

$$\mathcal{F}_n(\delta) = \mathcal{F}_n(2) = \{h = (f - g) : (f, g) \in \mathcal{F}_n\} \supset \mathcal{F}_n$$

Thus, using elementary manipulations we prove that the condition **(A2)** implies that

$$\|\Phi_{n+1}(\mu) - \Phi_{n+1}(\eta_n)\|_{\mathcal{F}_{n+1}} \leq c(n) \|\mu - \eta_n\|_{\Sigma_n(\mathcal{F}_{n+1})}$$

for some separable collection  $\Sigma_n(\mathcal{F}_{n+1})$  of measurable functions  $f_n$  on  $E_n$ , s.t.  $\|f_n\| \leq 1$ , and such that

$$I(\Sigma_n(\mathcal{F}_{n+1})) < c'(n) I(\mathcal{F}_{n+1}) \quad (5.1)$$

for some finite constants  $c(n)$  and  $c'(n) < \infty$ .

This implies that

$$\begin{aligned} \sqrt{N} \|\Phi_{n+1}(\eta_n^N) - \Phi_{n+1}(\eta_n)\|_{\mathcal{F}_{n+1}} &\leq c(n) \sqrt{N} \|\eta_n^N - \eta_n\|_{\Sigma_n(\mathcal{F}_{n+1})} \\ &= c(n) \|W_n^N\|_{\Sigma_n(\mathcal{F}_{n+1})} \end{aligned} \quad (5.2)$$

On the other hand, we have

$$W_{n+1}^N = V_{n+1}^N + \sqrt{N} [\Phi_{n+1}(\eta_n^N) - \Phi_{n+1}(\eta_n)]$$

and therefore

$$\begin{aligned} \|W_{n+1}^N\|_{\mathcal{F}_{n+1}} &\leq \|V_{n+1}^N\|_{\mathcal{F}_{n+1}} + c(n) \|W_n^N\|_{\Sigma_n(\mathcal{F}_{n+1})} \\ &\leq \sum_{p=0}^{n+1} c_p(n) \|V_p^N\|_{\Sigma_{p,n}(\mathcal{F}_{n+1})} \end{aligned}$$

with  $\Sigma_{p,n} = \Sigma_p \circ \Sigma_{p+1,n}$ , and  $c_p(n) = \prod_{p \leq q < n} c(q)$ . We let  $\pi_\psi[Y]$  be the Orlicz norm of an  $\mathbb{R}$ -valued random variable  $Y$  associated with the convex function  $\psi(u) = e^{u^2} - 1$ , and defined by

$$\pi_\psi(Y) = \inf \{a \in (0, \infty) : \mathbb{E}(\psi(|Y|/a)) \leq 1\}$$

with the convention  $\inf_\emptyset = \infty$ . From previous calculations, we have

$$\pi_\psi \left( \|W_{n+1}^N\|_{\mathcal{F}_{n+1}} \right) \leq \sum_{p=0}^{n+1} c_p(n) \pi_\psi \left( \|V_p^N\|_{\Sigma_{p,n}(\mathcal{F}_{n+1})} \right)$$

Combining Lemma 8.1 with (5.1), we find that

$$\pi_\psi \left( \|W_{n+1}^N\|_{\mathcal{F}_{n+1}} \right) \leq c''(n) I(\mathcal{F}_{n+1})$$

for some finite constants  $c''(n)$ . By (5.2), we also have that

$$\sqrt{N} \pi_\psi \left( \|\Phi_{n+1}(\eta_n^N) - \Phi_{n+1}(\eta_n)\|_{\mathcal{F}_{n+1}} \right) \leq c'''(n) I(\mathcal{F}_{n+1})$$



for some finite constants  $c'''(n)$ . This shows that the random fields  $V_n^N$  satisfy the regularity condition stated in (8.2).

Arguing as above, we prove that

$$\|W_n^N\|_{\mathcal{F}_n(\delta)} \leq \sum_{p=0}^n \alpha_p(n) \|V_p^N\|_{\mathcal{F}_{p,n}(\beta_p(n)\delta)}$$

for some separable collection  $\mathcal{F}_{p,n}$  of measurable functions  $f_p$  on  $E_p$ , s.t.  $\|f_p\| \leq 1$ , and such that  $I(\mathcal{F}_{p,n}) < \infty$ , and for some finite constants  $\alpha_p(n)$  and  $\beta_p(n) < \infty$ .

$$\mathbb{P}\left(\|W_n^N\|_{\mathcal{F}_n(\delta)} > y\sqrt{\alpha(N)}\right) \leq \sum_{p=0}^n \mathbb{P}\left(\|V_p^N\|_{\mathcal{G}_{p,n}(\delta)} > y_{p,n}\sqrt{\alpha(N)}\right)$$

with  $y_{p,n} = y/[(n+1)\alpha_p(n)]$  and  $\mathcal{G}_{p,n}(\delta) := \mathcal{F}_{p,n}(\beta_p(n)\delta)$ . On the other hand, using lemma 8.2, we have

$$\begin{aligned} & \frac{1}{\alpha(N)} \log \mathbb{P}\left(\|V_p^N\|_{\mathcal{G}_{p,n}(\delta)} > y_{p,n}\sqrt{\alpha(N)}\right) \\ & \leq -\frac{y_{p,n}^2}{2a_{(\beta_p(n)\delta)}(\mathcal{F}_{p,n})^2} \left(1 - \frac{\alpha(N)}{N} y_{p,n}^2 \left(\frac{b_{(\beta_p(n)\delta)}(\mathcal{F}_{p,n})}{a_{(\beta_p(n)\delta)}(\mathcal{F}_{p,n})}\right)^2\right) \xrightarrow{N \uparrow \infty} -\frac{y_{p,n}^2}{2a_{(\beta_p(n)\delta)}(\mathcal{F}_{p,n})^2} \end{aligned}$$

with some finite constant  $b_\delta(\mathcal{F})$ , and

$$a_\delta(\mathcal{F}) \leq c \int_0^\delta \sqrt{\log \mathcal{N}(\mathcal{F}, \epsilon)} d\epsilon \xrightarrow{\delta \downarrow 0} 0$$

so that

$$-\frac{y_{p,n}^2}{2a_{(\beta_p(n)\delta)}(\mathcal{F}_{p,n})^2} \xrightarrow{\delta \downarrow 0} -\infty$$

This ends the proof of the lemma. ■

## 6 Appendix A.

### 6.1 Proof of lemma 3.2

The first  $\mathbb{L}_m$  almost sure estimates is a direct consequence of Kintchine's inequality, let us examine some direct consequences of this result. Combining the Lipschitz property  $(\Phi_{p,n})$  of the semigroup  $\Phi_{p,n}$  with the decomposition

$$[\eta_n^N - \eta_n] = \sum_{p=0}^n [\Phi_{p,n}(\eta_p^N) - \Phi_{p,n}(\Phi_p(\eta_{p-1}^N))]$$

we find that (by condition (K))

$$\sqrt{N} |[\eta_n^N - \eta_n](f_n)| = \sum_{p=0}^n \int |V_p^N(h)| T_{\Phi_p(\eta_{p-1}^N)}^{\Phi_{p,n}}(f, dh)$$

In the above displayed formulae, we have used the convention  $\Phi_0(\eta_{-1}^N) = \eta_0$ , for  $p = 0$ . The proof of (3.3) is a direct consequence of the previous  $\mathbb{L}_m$  almost sure estimates. On the other hand, using decomposition

$$W_n^N = \sqrt{N} \sum_{p=0}^n [\Phi_{p,n}(\eta_p^N) - \Phi_{p,n}(\Phi_p(\eta_{p-1}^N))]$$

we find that  $W_n^N = I_n^N + J_n^N$ , with the pair of random measures  $(I_n^N, J_n^N)$  given by

$$I_n^N := \sum_{p=0}^n V_p^N \mathcal{D}_{p,n}^{(N)} \quad \text{and} \quad J_n^N := \sqrt{N} \sum_{p=0}^n \mathcal{R}_{p,n}(\eta_p^N, \Phi_p(\eta_{p-1}^N))$$

with

$$\mathcal{D}_{p,n}^{(N)} := \mathcal{D}_{\Phi_p(\eta_{p-1}^N)} \Phi_{p,n} \quad \text{and} \quad \mathcal{R}_{p,n} = \mathcal{R}^{\Phi_{p,n}}.$$

Under our assumptions, we have the almost sure estimates

$$\sup_{N \geq 1} \beta(\mathcal{D}_{p,n}^{(N)}) \leq \beta(\mathcal{D}\Phi_{p,n}) := \sup_{\eta \in \mathcal{P}(E_p)} \beta(\mathcal{D}_\eta \Phi_{p,n}).$$

Using the generalized Minkowski integral inequality we find that

$$N \mathbb{E} \left( \left| \mathcal{R}_{p,n}(\eta_p^N, \Phi_p(\eta_{p-1}^N))(f_n) \right| \mid \mathcal{A}_{p-1}^{(N)} \right) \leq \delta(R^{\Phi_{p,n}})$$

from which we readily conclude that

$$\mathbb{E} \left( \left| \sqrt{N} J_n^N(f_n) \right| \right) = N \mathbb{E} \left( \left| \sum_{p=0}^n \mathcal{R}_{p,n}(\eta_p^N, \Phi_p(\eta_{p-1}^N))(f_n) \right| \right) \leq \sum_{p=0}^n \delta(R^{\Phi_{p,n}})$$

The proof of (3.4) is now clear. This ends the proof of the lemma. ■

## 6.2 Proof of lemma 3.5

Using Holder inequality, for any  $\delta > 0$  we find that

$$\mathbb{E} \left( e^{\alpha(N) B^N(t)} \right) \leq \mathbb{E} \left( e^{\alpha(N) (1+\delta)A^N(t)} \right)^{\frac{1}{1+\delta}} \mathbb{E} \left( e^{\alpha(N) \frac{1+\delta}{\delta} |[A^N - B^N](t)|} \right)^{\frac{\delta}{1+\delta}}.$$

Under our assumptions, we have

$$|(1+\delta)A^N(t) - A^N((1+\delta)t)| \leq (1+\delta) a_t(\delta) \quad \text{with} \quad \lim_{\delta \rightarrow 0} a_t(\delta) = 0$$

this implies that

$$\mathbb{E} \left( e^{\alpha(N) (1+\delta)A^N(t)} \right)^{\frac{1}{1+\delta}} \leq \mathbb{E} \left( e^{\alpha(N) A^N((1+\delta)t)} \right)^{\frac{1}{1+\delta}} e^{a_t(\delta)}$$

and therefore

$$\limsup_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N)(1+\delta)A^N(t)} \right)^{\frac{1}{1+\delta}} \leq \frac{1}{1+\delta} \Lambda((1+\delta)t) + a_t(\delta).$$

One conclude that

$$\limsup_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N)B^N(t)} \right) \leq \frac{1}{1+\delta} \Lambda((1+\delta)t) + a_t(\delta) \xrightarrow{\delta \rightarrow 0} \Lambda(t).$$

In much the same way, if we set  $t^\delta = t/(1+\delta)$  we have

$$\mathbb{E} \left( e^{\alpha(N)A^N(t^\delta)} \right) \leq \mathbb{E} \left( e^{\alpha(N)(1+\delta)B^N(t^\delta)} \right)^{\frac{1}{1+\delta}} \mathbb{E} \left( e^{\alpha(N)\frac{1+\delta}{\delta}|[B^N - A^N](t^\delta)|} \right)^{\frac{\delta}{1+\delta}}.$$

Under our assumptions, we have

$$|(1+\delta)B^N(t/(1+\delta)) - B^N(t)| \leq b_t(\delta) \quad \text{with} \quad \lim_{\delta \rightarrow 0} b_t(\delta) = 0$$

this implies that

$$\mathbb{E} \left( e^{\alpha(N)(1+\delta)B^N(t^\delta)} \right) \leq \mathbb{E} \left( e^{\alpha(N)B^N(t)} \right) e^{b_t(\delta)}$$

and therefore

$$\liminf_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N)(1+\delta)B^N(t^\delta)} \right) \leq \liminf_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N)B^N(t)} \right) + b_t(\delta).$$

One conclude that

$$(1+\delta)\Lambda(t/(1+\delta)) \leq \liminf_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N)B^N(t)} \right) + b_t(\delta)$$

and letting  $\delta \downarrow 0$  we find that

$$\Lambda(t) \leq \liminf_{N \rightarrow \infty} \frac{1}{\alpha(N)} \log \mathbb{E} \left( e^{\alpha(N)B^N(t)} \right).$$

This ends the proof of the lemma. ■

### 6.3 Proof of lemma 3.7

Taking the logarithm, we find that

$$\log \mathcal{Z}_n^{(N)}(f) = \sum_{p=0}^n \Delta_p \log (\mathcal{Z}^{(N)}(f))$$

with the random increments

$$\Delta_n \log (\mathcal{Z}^{(N)}(f)) = \log \mathbb{E} \left( \exp \left\{ \sqrt{\alpha(N)} V_n^N(f_n) \right\} \mid \mathcal{A}_{n-1}^{(N)} \right)$$

We observe that

$$\Delta_n \log (\mathcal{Z}^N(f)) = \sum_{i=1}^N \log \mathbb{E} \left( \exp \left\{ X_n^{(N,i)}(f_n) \right\} \mid \mathcal{A}_{n-1}^{(N)} \right)$$

with the sequence of random variables

$$X_n^{(N,i)}(f_n) = \sqrt{\frac{\alpha(N)}{N}} \left( f_n(\xi_n^{N,i}) - K_{n,\eta_{n-1}^N}(f_n)(\xi_{n-1}^{N,i}) \right)$$

such that

$$\begin{aligned} \mathbb{E} \left( X_n^{(N,i)}(f_n) \mid \mathcal{A}_{n-1}^{(N)} \right) &= 0 \\ \mathbb{E} \left( X_n^{(N,i)}(f_n)^2 \mid \mathcal{A}_{n-1}^{(N)} \right) &= \frac{\alpha(N)}{N} K_{n,\eta_{n-1}^N} \left[ \left( f_n - K_{n,\eta_{n-1}^N}(f_n) \right)^2 \right] (\xi_{n-1}^{N,i}) \end{aligned}$$

We recall that for every centered random variable  $X$  with  $|X| \leq c$  for some  $c < \infty$ , we have

$$-\frac{\sigma^2}{2} \epsilon_1(c) \leq \log \mathbb{E}(e^X) - \frac{\sigma^2}{2} \leq \frac{\sigma^2}{2} \epsilon_2(c)$$

with the parameters  $(\sigma^2, \epsilon_1(c), \epsilon_2(c))$  given below

$$\sigma^2 = \mathbb{E}(X^2) \quad \epsilon_1(c) := \left\{ [1 - \theta(-c)] + \left( \frac{\theta(-c) c}{2} \right)^2 \right\} \quad \text{and} \quad \epsilon_2(c) := [\theta(c) - 1].$$

In the above display,  $\theta$  is the  $\mathcal{C}^1$ -increasing function defined by  $\theta(x) = \frac{2}{x^2} (e^x - 1 - x)$  for  $x \neq 0$  and  $\theta(0) = 1$ . We set

$$\forall j = 1, 2 \quad \tau_{j,n}^{(N)}(f) := \sup_{0 \leq p \leq n} \epsilon_j \left( \sqrt{\frac{\alpha(N)}{N}} \text{osc}(f_p) \right).$$

Using the above estimate, for any  $p \leq n$  we find that

$$\Delta_p \log (\mathcal{Z}^{(N)}(f)) - \frac{\alpha(N)}{2} \Delta_n \langle M^{(N)}(f) \rangle \leq \tau_{2,n}^{(N)}(f) \frac{\alpha(N)}{2} \Delta_p \langle M^{(N)}(f) \rangle$$

and

$$-\tau_{1,n}^{(N)}(f) \frac{\alpha(N)}{2} \Delta_p \langle M^{(N)}(f) \rangle \leq \Delta_p \log (\mathcal{Z}^N(f)) - \frac{\alpha(N)}{2} \Delta_p \langle M^{(N)}(f) \rangle.$$

This yields that

$$-\tau_{1,n}^{(N)}(f) \frac{\alpha(N)}{2} \langle M^{(N)}(f) \rangle_n \leq \log \mathcal{Z}_n^{(N)}(f) - \frac{\alpha(N)}{2} \langle M^{(N)}(f) \rangle_n \leq \tau_{2,n}^{(N)}(f) \frac{\alpha(N)}{2} \langle M^{(N)}(f) \rangle_n.$$

The end of the proof is now a direct consequence of the following formula

$$\begin{aligned} &\exp \left\{ \sqrt{\alpha(N)} M_n^{(N)}(f) - \frac{\alpha(N)}{2} \langle M^{(N)}(f) \rangle_n \right\} \\ &= E_n^{(N)}(f) \exp \left\{ \log \mathcal{Z}_n^{(N)}(f) - \frac{\alpha(N)}{2} \langle M^{(N)}(f) \rangle_n \right\} \end{aligned}$$

This ends the proof of the lemma. ■

## 7 Appendix B.

Given two measurable spaces  $E_1$  and  $E_2$ , we consider a probability measure  $\mu$  over the set  $E_1$  and a Markov transition  $M(x, dy)$  from  $E_1$  to  $E_2$ . We let  $V$  be the Gaussian and centered random field on  $\mathbb{L}_2(E_2, \mu M)$  defined for any  $f \in \mathbb{L}_2(E_2, \mu M) := \mathbb{L}_2(\mu M)$  by

$$\mathbb{E}(V(f)^2) = \mu(M([f - M(f)]^2)). \quad (7.1)$$

In the above display, we have used the convention

$$x \mapsto M([f - M(f)]^2)(x) = M(f^2)(x) - M(f)(x)^2 \quad \text{and} \quad M(f)(x) = \int M(x, dy) f(y).$$

**Definition 7.1** For any  $w \in \mathcal{M}(E_2)$  we set

$$I(w) = \sup_{f \in \mathcal{B}(E)} \left( w(f) - \frac{1}{2} \mathbb{E}(V(f)^2) \right). \quad (7.2)$$

**Lemma 7.2** Assume that  $E_1 = E_2$  and  $M(x, dy) = \mu(dy)$ . In this situation, we have

$$I(w) = \sup_{f \in \mathbb{L}_2(\mu)} \left( w(f) - \frac{1}{2} \mu([f - \mu(f)]^2) \right) \quad (7.3)$$

and

$$I(w) = I_\mu(w) := \begin{cases} \frac{1}{2} \left\| \frac{dw}{d\mu} \right\|_{\mathbb{L}_2(\mu)}^2, & \text{if } w \ll \mu \text{ with } \frac{dw}{d\mu} \in \mathbb{L}_2(\mu) \text{ and } w(E) = 0 \\ +\infty, & \text{otherwise.} \end{cases} \quad (7.4)$$

**Proof:**

To check this claim, firstly we notice that for any constant function  $f(x) = a$  we have

$$w(f) - \mu([f - \mu(f)]^2) = a w(E).$$

Choosing  $a = b w(E)$  with  $b \in \mathbb{R}_+$ , we readily find that

$$w(E) \neq 0 \Rightarrow \forall b \in \mathbb{R}_+ \quad I(w) \geq b w(E)^2 \Rightarrow I(w) = \infty.$$

Whenever  $w$  is not absolutely continuous w.r.t.  $\mu$ , we can find a measurable set  $A \in \mathcal{E}$  such that  $w(A) \neq 0$  and  $\mu(A) = 0$ . In this situation, we have

$$\forall a \in \mathbb{R}_+ \quad f = a w(A) 1_A \Rightarrow w(f) - \mu([f - \mu(f)]^2) = a w(A)^2 \Rightarrow I(w) = \infty.$$

On the other hand, using the fact that

$$f = g + a \Rightarrow w(f) - \frac{1}{2} \mu([f - \mu(f)]^2) = w(g) - \frac{1}{2} \mu([g - \mu(g)]^2)$$

as soon as  $w(E) = 0$ , we can reduce the supremum in (7.3) to functions  $f$  with  $\mu(f) = 0$ . This yields

$$I(w) = \sup_{f \in \mathbb{L}_{0,2}(\mu)} \left( \mu \left( \frac{dw}{d\mu} f \right) - \frac{1}{2} \mu(f^2) \right) \quad \text{with} \quad \mathbb{L}_{0,2}(\mu) = \{h \in \mathbb{L}_2(\mu) : \mu(h) = 0\}.$$

Finally, we observe that

$$\mu \left( \frac{dw}{d\mu} f \right) - \frac{1}{2} \mu(f^2) = \frac{1}{2} \mu \left( \left( f - \frac{dw}{d\mu} \right)^2 \right) + \mu \left[ \left( \frac{dw}{d\mu} \right)^2 \right].$$

Choosing  $f = \frac{dw}{d\mu}$  we prove (7.3). This ends the proof of the lemma.  $\blacksquare$

The analysis of the variational formula (7.2) for more general Markov transitions  $M$  is a little more involved. Before getting into further details, we observe that

$$\begin{aligned} \mu(M([f - M(f)]^2)) &= (\mu M) \left( ([f - (\mu M)(f)]^2) \right) - \mu \left( ([M(f) - \mu M(f)]^2) \right) \\ &\leq (\mu M) \left( ([f - (\mu M)(f)]^2) \right). \end{aligned}$$

The above inequality implies that

$$I(w) := \sup_{f \in \mathcal{B}(E)} \left( w(f) - \frac{1}{2} \mu(M([f - M(f)]^2)) \right) \geq I_{\mu M}(w) \quad (7.5)$$

where  $I_{\mu M}(w)$  is given by (7.4) and therefore

$$I(w) < \infty \Rightarrow w \ll (\mu M) \quad \text{with} \quad \frac{dw}{d\mu M} \in \mathbb{L}_2(\mu M) \quad \text{and} \quad w(E) = 0.$$

Next, we follow the analysis developed in [19]. Firstly, we notice that  $M$  is an operator  $\mathbb{L}_2(\mu M)$  into  $\mathbb{L}_2(\mu)$ .

**Definition 7.3** We let  $M_\mu^*$  be the adjoint operator of  $M$  from  $\mathbb{L}_2(\mu)$  into  $\mathbb{L}_2(\mu M)$  given by

$$\forall (f, g) \in \mathbb{L}_2(\mu) \times \mathbb{L}_2(\mu) \quad (\mu M) (f M_\mu^*(g)) = \mu(M(f) g).$$

In fact  $M_\mu^*$  can be identified as a kernel  $M_\mu^*(x_2, dx_1)$  which is the conditional law of  $x_1$  knowing  $x_2$  under the probability measure  $\mu \otimes M(dx_1, dx_2) := \mu(dx_1)M(x_1, dx_2)$  on  $E_1 \times E_2$ . By construction, we have

$$\mu(M([f - M(f)]^2)) = \mu(M(f^2) - M(f)^2) = (\mu M) (f (Id - M_\mu^* M) f).$$

We observe that  $(M_\mu^* M)$  is a self adjoint operator on  $\mathbb{L}_2(\mu M)$  with

$$(\mu M) = (\mu M)(M_\mu^* M).$$

Also notice that

$$(\mu M) (f (Id - M_\mu^* M) f) = \int \mu(dx) M(x, dy) [f(y) - M(f)(x)]^2$$

from which, we see that

$$(\mu M) (f (Id - M_\mu^* M) f) = 0 \Leftrightarrow f(y) = M f(x), \quad \mu(dx) M(x, dy) - a.s. \quad (7.6)$$

Let  $\mathcal{N}$  be the subspace of those elements  $h$  in  $\mathbb{L}_2(E_2, \mu M)$  such that  $h(y) = M h(x)$ ,  $\mu(dx) M(x, dy) - a.s.$ . Notice that  $\mathcal{N} = \{h \in L^2(E_2, \mu M); (Id - M_\mu^* M)h = 0\}$ . It is well known that  $\mathcal{N} = \mathbb{L}_2(E_2, \mathcal{G}, \mu M)$  where  $\mathcal{G}$  is the sub- $\sigma$ -field generated by all  $h \in \mathcal{N}$  ([16]). In particular  $\mathcal{N} \cap \mathbb{L}_\infty(E_2, \mu M)$  is dense in  $\mathcal{N}$ .

Consider the orthogonal supplementary subspace  $\mathcal{H}_0(\mu M)$  of  $\mathcal{N}$  in  $\mathbb{L}_2(\mu M)$ .

In this notation, the rate function  $I$  defined in (7.2) takes the form

$$I(w) = \sup_{f \in \mathcal{H}_0(\mu M)} \left( w(f) - \frac{1}{2} (\mu M) (f (Id - M_\mu^* M) f) \right).$$

Before getting into further details, arguing as in the proof of lemma 7.2, we notice that

$$I(w) < \infty \Rightarrow w \ll \mu M \quad \text{and} \quad w(h) = 0, \quad \forall h \in \mathcal{N} \cap \mathbb{L}_\infty(\mu M).$$

As  $M_\mu^* M$  is self-adjoint, definite nonnegative on  $\mathbb{L}_2(\mu M)$  and its norm is 1, we can write the spectral decomposition of  $(Id - M_\mu^* M)$  on  $\mathcal{H}_0(\mu M)$

$$(Id - M_\mu^* M) = \int_{[0,1]} \lambda dE_\lambda = \int_{(0,1]} \lambda dE_\lambda$$

( $E_0 = 0$  by the very definition of  $\mathcal{H}_0(\mu M)$ ). The operator  $(Id - M_\mu^* M) : \mathcal{H}_0(\mu M) \rightarrow \mathcal{H}_0(\mu M)$  is injective and its inverse is given by

$$R_{0,\mu} = (Id - M_\mu^* M)^{-1} : \text{Dom}(R_{0,\mu}) \subset \mathcal{H}_0(\mu M) \mapsto \mathcal{H}_0(\mu M)$$

with

$$R_{0,\mu} = \int_{(0,1]} \frac{1}{\lambda} dE_\lambda \quad \text{and} \quad \text{Dom}(R_{0,\mu}) := \left\{ h \in \mathcal{H}_0(\mu M) : \int_{(0,1]} \frac{1}{\lambda^2} d\langle E_\lambda(h), h \rangle < \infty \right\}.$$

**Definition 7.4** We let  $\mathcal{H}_1(\mu M)$  be the completion of the pre-Hilbert space  $\mathcal{H}_0(\mu M)$  with the inner product given by

$$\begin{aligned} \langle f, g \rangle_1 &:= \langle f, (Id - M_\mu^* M)(g) \rangle \\ &= \mu M (f (Id - M_\mu^* M)(g)) = \int_{(0,1]} \lambda d\langle E_\lambda(f), g \rangle \end{aligned}$$

We define  $(\mathcal{H}_{-1}(\mu M), \|\cdot\|_{-1})$  as the dual space of  $(\mathcal{H}_1(\mu M), \|\cdot\|_1)$  w.r.t. the canonical dual relation  $\mathcal{H}_0(\mu M) = \mathcal{H}_0(\mu M)'$ .

By [19],  $\mathcal{H}_{-1}(E_2, \mu M)$  is the subspace of  $f \in \mathcal{H}_0(\mu M)$  such that  $\|f\|_{-1} < +\infty$ , and  $R_{0,\mu}$  can be regarded as an isomorphism from  $\mathcal{H}_{-1}(\mu M)$  to  $\mathcal{H}_1(\mu M)$ ; furthermore for any  $f \in \mathcal{H}_0(E_2, \mu M)$ ,

$$\|f\|_{-1}^2 = \int_{(0,1]} \frac{1}{\lambda} d\langle E_\lambda(f), f \rangle = \sum_{n=0}^{\infty} \langle f, (M_\mu^* M)^n f \rangle. \quad (7.7)$$

Notice also that given  $f \in \mathbb{L}_2(\nu M)$  if  $\sum_{n=0}^{\infty} \langle f, (M_{\mu}^* M)^n f \rangle < +\infty$ , then  $f \in \mathcal{H}_0(E_2, \mu M)$ .

We further assume that  $w \ll \mu M$ , and  $h_w = \frac{dw}{d\mu M}$  satisfies  $\langle h_w, h \rangle = 0$  for all  $h \in \mathcal{N} \cap \mathbb{L}_{\infty}(\mu M)$ . In this situation, if

$$I(w) = \sup_{f \in \mathcal{H}_1(\mu M) \cap \mathbb{L}_{\infty}(\mu M)} \left( \langle h_w, f \rangle - \frac{1}{2} \|f\|_1^2 \right) < \infty$$

then  $f \rightarrow \langle h_w, f \rangle$  is a bounded linear form on  $\mathcal{H}_1(\mu M) \cap \mathbb{L}_{\infty}(\mu M)$  w.r.t. the norm  $\|\cdot\|_1$ . This yields that  $h_w \in \mathcal{H}_{-1}(\mu M)$  and

$$\begin{aligned} \langle h_w, f \rangle - \frac{1}{2} \|f\|_1^2 &= \langle R_{0,\mu}(h_w), (Id - M_{\mu}^* M)(f) \rangle - \frac{1}{2} \|f\|_1^2 \\ &= \langle R_{0,\mu}(h_w), f \rangle_1 - \frac{1}{2} \|f\|_1^2 \\ &= -\frac{1}{2} \|f - R_{0,\mu}(h_w)\|_1^2 + \frac{1}{2} \langle R_{0,\mu}(h_w), h_w \rangle \end{aligned}$$

from which we conclude that

$$I(w) = \frac{1}{2} \langle R_{0,\mu}(h_w), h_w \rangle = \frac{1}{2} \|h_w\|_{-1}^2 = \frac{1}{2} \sum_{n=0}^{\infty} \langle h_w, (M_{\mu}^* M)^n h_w \rangle.$$

In summary we have proven

**Proposition 7.5** *The rate function defined in (7.2) is given by*

$$I(w) = \begin{cases} \frac{1}{2} \sum_{n=0}^{\infty} \langle h_w, (M_{\mu}^* M)^n h_w \rangle, & \text{if } w \ll \mu M, h_w = \frac{dw}{d(\mu M)} \in \mathbb{L}_2(\mu M) \\ +\infty, & \text{otherwise.} \end{cases}$$

## 8 Appendix C.

In the further development of this section,  $c < \infty$  stands for some finite universal constant, whose values may vary from line to line.

Let  $(\mu^i)_{i \geq 1}$  be a sequence of probability measures on a given measurable state space  $(E, \mathcal{E})$ . During the further development of this section, we fix an integer  $N \geq 1$ . To clarify the presentation, we slightly abuse the notation and we denote respectively by

$$m(X) = \frac{1}{N} \sum_{i=1}^N \delta_{X^i} \quad \text{and} \quad \mu = \frac{1}{N} \sum_{i=1}^N \mu^i$$

the  $N$ -empirical measure associated with a collection of independent random variables  $X = (X^i)_{i \geq 1}$ , with respective distributions  $(\mu^i)_{i \geq 1}$ , and the  $N$ -averaged measure associated with the sequence of measures  $(\mu^i)_{i \geq 1}$ . We also consider the empirical random field sequences

$$V(X) = \sqrt{N} (m(X) - \mu)$$

We also set

$$\sigma(f)^2 := \mathbb{E} (V(X)(f)^2) = \frac{1}{N} \sum_{i=1}^N \mu^i ([f - \mu^i(f)]^2) \quad (8.1)$$



Let  $\mathcal{F}$  be a given collection of measurable functions  $f : E \rightarrow \mathbb{R}$  such that  $\|f\| \leq 1$ . No generality is lost and much convenience is gained by supposing that the unit and the null functions  $f = \mathbb{1}$  and  $f = 0 \in \mathcal{F}$ . Furthermore, to avoid some unnecessary technical measurability questions, we shall also suppose that  $\mathcal{F}$  is separable in the sense that it contains a countable and dense subset. For any separable class of uniformly bounded functions  $\mathcal{H}$  s.t.  $\sup_{h \in \mathcal{H}} \|h\| \leq H$  we set

$$I(\mathcal{H}) = \int_0^{2H} \sqrt{\log \mathcal{N}(\mathcal{H}, \epsilon)} d\epsilon$$

We further assume that there exists some probability measure  $\bar{\mu}$  on  $E$  such that

$$\sqrt{N} \pi_\psi (\|\mu - \bar{\mu}\|_{\mathcal{F}}) \leq \tau(I(\mathcal{F})) \quad (8.2)$$

for any class of function  $\mathcal{F}$  satisfying the above properties, with finite entropy  $I(\mathcal{F}) < \infty$ , and some non decreasing function  $\tau$ . In the above displayed formula,  $\pi_\psi[Y]$  stands for the Orlicz norm of an  $\mathbb{R}$ -valued random variable  $Y$  associated with the convex function  $\psi(u) = e^{u^2} - 1$ , and defined by

$$\pi_\psi(Y) = \inf \{a \in (0, \infty) : \mathbb{E}(\psi(|Y|/a)) \leq 1\}$$

with the convention  $\inf_{\emptyset} = \infty$ . We recall that

$$\mathbb{E}(e^{tY}) \leq 2 \exp\left(\frac{t^2}{4} \pi_\psi(Y)^2\right) \quad (8.3)$$

for any  $t \geq 0$ . We prove this claim using the estimate

$$tY = \left(\frac{t\pi_\psi(Y)}{\sqrt{2}}\right) \left(\frac{\sqrt{2} Y}{\pi_\psi(Y)}\right) \leq \frac{(t\pi_\psi(Y))^2}{4} + \left(\frac{Y}{\pi_\psi(Y)}\right)^2$$

We consider the possibly bias random field sequence

$$\bar{V}(X) = \sqrt{N} (m(X) - \bar{\mu}) = V(X) + \sqrt{N} (\mu - \bar{\mu})$$

The following lemma is satisfied without the regularity condition (8.2).

**Lemma 8.1**

$$\pi_\psi (\|V(X)\|_{\mathcal{F}}) \leq c I(\mathcal{F})$$

**Proof:**

We consider a collection of independent copies  $X' = (X'^i)_{i \geq 1}$  of the random variables  $X = (X^i)_{i \geq 1}$ . Let  $\epsilon = (\epsilon_i)_{i \geq 1}$  constitute a sequence that is independent and identically distributed with

$$P(\epsilon_1 = +1) = P(\epsilon_1 = -1) = 1/2$$

We also consider the empirical random field sequences

$$V_\epsilon(X) := \sqrt{N} m_\epsilon(X)$$

We also assume that  $(\epsilon, X, X')$  are independent. We associate with the pairs  $(\epsilon, X)$  and  $(\epsilon, X')$  the random measures  $m_\epsilon(X) = \frac{1}{N} \sum_{i=1}^N \epsilon_i \delta_{X^i}$  and  $m_\epsilon(X') = \frac{1}{N} \sum_{i=1}^N \epsilon_i \delta_{X'^i}$ .

We notice that

$$\begin{aligned} \|m(X) - \mu\|_{\mathcal{F}}^p &= \sup_{f \in \mathcal{F}} |m(X)(f) - \mathbb{E}(m(X')(f))|^p \\ &\leq \mathbb{E}(\|m(X) - m(X')\|_{\mathcal{F}}^p | X) \end{aligned}$$

and in view of the symmetry of the random variables  $(f(X^i) - f(X'^i))_{i \geq 1}$  we have

$$\mathbb{E}(\|m(X) - m(X')\|_{\mathcal{F}}^p) = \mathbb{E}(\|m_\epsilon(X) - m_\epsilon(X')\|_{\mathcal{F}}^p)$$

from which we conclude that

$$E(\|V(X)\|_{\mathcal{F}}^p) \leq 2^p E(\|V_\epsilon(X)\|_{\mathcal{F}}^p) \quad (8.4)$$

By using the Chernov-Hoeffding inequality for any  $x = (x^1, \dots, x^N) \in E^N$ , the empirical process

$$f \longrightarrow V_\epsilon(x)(f) := \sqrt{N} m_\epsilon(x)(f)$$

is sub-Gaussian for the norm  $\|f\|_{L_2(m(x))} = m(x)(f^2)^{1/2}$ . Namely, for any couple of functions  $f, g$  and any  $\delta > 0$  we have

$$\mathbb{E}\left([V_\epsilon(x)(f) - V_\epsilon(x)(g)]^2\right) = \|f - g\|_{\mathbb{L}_2(m(x))}^2$$

and by Hoeffding's inequality

$$P(|V_\epsilon(x)(f) - V_\epsilon(x)(g)| \geq \delta) \leq 2 e^{-\frac{1}{2}\delta^2/\|f-g\|_{\mathbb{L}_2(m(x))}^2}$$

If we set  $Z = \left(\frac{V_\epsilon(x)(f)}{\sqrt{6}\|f\|_{\mathbb{L}_2(m(x))}}\right)^2$ , then we find that

$$\begin{aligned} \mathbb{E}(e^Z) - 1 &= \int_0^\infty e^t \mathbb{P}(Z \geq t) dt \\ &= \int_0^\infty e^t \mathbb{P}\left(|V_\epsilon(x)(f)| \geq \sqrt{6t} \|f\|_{\mathbb{L}_2(m(x))}\right) dt \\ &\leq 2 \int_0^\infty e^t e^{-3t} dt = 1 \end{aligned}$$

from which we conclude that

$$\pi_\psi(V_\epsilon(x)(f) - V_\epsilon(x)(g)) \leq \sqrt{6}\|f - g\|_{\mathbb{L}_2(m(x))}$$

By the maximal inequalities for sub Gaussian processes (see [6], [20]), we find that

$$\pi_\psi(\|V_\epsilon(x)\|_{\mathcal{F}}) \leq c I(\mathcal{F}) < \infty$$

for any  $x \in E^N$ . By (8.4), this clearly implies that

$$\pi_\psi(\|V(X)\|_{\mathcal{F}}) \leq c I(\mathcal{F})$$

This ends the proof of the lemma. ■

For any  $\delta > 0$ , we also set

$$\mathcal{F}(\delta) := \left\{ h = (f - g) : (f, g) \in \mathcal{F} \text{ s.t. } \overline{\mu}(h^2)^{1/2} \leq \delta \right\}$$

Notice that

$$\mathcal{N}(\mathcal{F}(\delta), \mathbb{L}_2(\eta), \epsilon) \leq \mathcal{N}(\mathcal{F}(\infty), \mathbb{L}_2(\eta), \epsilon) \leq \mathcal{N}(\mathcal{F}, \mathbb{L}_2(\eta), \epsilon/2)^2$$

from which we conclude that

$$\mathcal{N}(\mathcal{F}(\delta), \epsilon) \leq \mathcal{N}(\mathcal{F}, \epsilon/2)^2$$

**Lemma 8.2** *Under the regularity condition (8.2), we have the following Laplace estimates*

$$\mathbb{E} \left( e^{t \|V(X)\|_{\mathcal{F}(\delta)}} \right) \leq 4 \exp \left( \frac{t^2}{2} \left[ a_\delta(\mathcal{F})^2 + \frac{1}{N} (tb_\delta(\mathcal{F})) \right] \right)$$

for any  $t \geq 0$ , with the parameters

$$a_\delta(\mathcal{F}) \leq c \int_0^\delta \sqrt{\log \mathcal{N}(\mathcal{F}, \epsilon)} d\epsilon$$

and

$$b_\delta(\mathcal{F}) \leq c \log \mathcal{N}(\mathcal{F}, \delta) [I(\mathcal{F}) + \tau(c I(\mathcal{F}))]$$

On the other hand, for any  $\delta > 0$  and any  $x \geq 0$ , we have

$$\log \mathbb{P} \left( \|V(X)\|_{\mathcal{F}(\delta)} \geq x \right) \leq - \sup_{t \geq 0} \left( tx - \frac{t^2}{2} \left[ a_\delta(\mathcal{F})^2 + \frac{1}{N} (tb_\delta(\mathcal{F}))^2 \right] \right)$$

Explicit calculations of the Legendre-Fenchel transformation can be derived, by choosing  $t = x/\alpha_{\mathcal{F}(\delta)}$ , we find the crude exponential concentration estimates

$$\frac{1}{\alpha(N)} \log \mathbb{P} \left( \frac{1}{\sqrt{\alpha(N)}} \|V(X)\|_{\mathcal{F}(\delta)} \geq x \right) \leq - \frac{x^2}{2a_\delta(\mathcal{F})^2} \left( 1 - \frac{\alpha(N)}{N} x^2 \left( \frac{b_\delta(\mathcal{F})}{a_\delta(\mathcal{F})} \right)^2 \right) \quad (8.5)$$

**Proof:**

For any probability measure  $\nu$ , we set

$$d_{2,\nu}(\mathcal{F}(\delta)) := \sup_{(h_1, h_2) \in \mathcal{F}(\delta)} \|h_1 - h_2\|_{\mathbb{L}_2(\nu)}$$

By definition, we clearly have that

$$d_{2,\mu}(\mathcal{F}(\delta)) \leq 2 \sup_{h \in \mathcal{F}(\delta)} \|h\|_{\mathbb{L}_2(\mu)} \leq 2\delta \quad (8.6)$$

Notice that for any couple of probability measures  $\nu_1, \nu_2$ , we have

$$d_{2,\nu_1}(\mathcal{F}(\delta)) \leq d_{2,\nu_2}(\mathcal{F}(\delta)) + \|\nu_1 - \nu_2\|_{\mathcal{G}(\delta)}^{1/2} \quad (8.7)$$

with

$$\mathcal{G}(\delta) = \{g = (h_1 - h_2)^2 : (h_1, h_2) \in \mathcal{F}(\delta)\}$$

By the maximal inequalities for sub Gaussian processes, we have the estimate

$$\pi_\psi(\|V_\epsilon(x)\|_{\mathcal{F}(\delta)}) \leq c \int_0^{d_{2,m(x)}(\mathcal{F}(\delta))} \sqrt{\log \mathcal{N}(\mathcal{F}(\delta), \epsilon)} d\epsilon$$

On the other hand, using (8.6) and (8.7) we prove that the r.h.s. integral is bounded by

$$\begin{aligned} & \int_0^{2\delta} \sqrt{\log \mathcal{N}(\mathcal{F}(\delta), \epsilon)} d\epsilon \\ & + \int_{2\delta}^{2\delta + \|m(x) - \bar{\mu}\|_{\mathcal{G}(\delta)}^{1/2}} \sqrt{\log \mathcal{N}(\mathcal{F}(\delta), \epsilon)} d\epsilon \\ & \leq \int_0^{2\delta} \sqrt{\log \mathcal{N}(\mathcal{F}(\delta), \epsilon)} d\epsilon + \sqrt{\log \mathcal{N}(\mathcal{F}(\delta), 2\delta)} \times \|m(x) - \bar{\mu}\|_{\mathcal{G}(\delta)}^{1/2} \\ & \leq c \int_0^{2\delta} \sqrt{\log \mathcal{N}(\mathcal{F}, \epsilon/2)} d\epsilon + c \sqrt{\log \mathcal{N}(\mathcal{F}, \delta)} \times \|m(x) - \bar{\mu}\|_{\mathcal{G}(\delta)}^{1/2} \end{aligned}$$

We conclude that

$$\pi_\psi(\|V_\epsilon(x)\|_{\mathcal{F}(\delta)}) \leq J_\delta(\mathcal{F}) + r_\delta(\mathcal{F}) \|m(x) - \bar{\mu}\|_{\mathcal{G}(\delta)}^{1/2}$$

with

$$J_\delta(\mathcal{F}) \leq c \int_0^\delta \sqrt{\log \mathcal{N}(\mathcal{F}, \epsilon)} d\epsilon \quad \text{and} \quad r_\delta(\mathcal{F}) \leq c \sqrt{\log \mathcal{N}(\mathcal{F}, \delta)}$$

Using (8.3), we have

$$\begin{aligned} \mathbb{E} \left( e^{t\|V_\epsilon(X)\|_{\mathcal{F}(\delta)}} \right) & \leq 2 \mathbb{E} \left[ e^{\frac{t^2}{2} (J_\delta(\mathcal{F})^2 + r_\delta(\mathcal{F})^2 \|m(X) - \bar{\mu}\|_{\mathcal{G}(\delta)})} \right] \\ & = 2 e^{\frac{t^2}{2} J_\delta(\mathcal{F})^2} \mathbb{E} \left[ e^{\frac{t^2}{2\sqrt{N}} r_\delta(\mathcal{F})^2 \|\bar{V}(X)\|_{\mathcal{G}(\delta)}} \right] \end{aligned}$$

and by (8.4) we have

$$\begin{aligned} \mathbb{E} \left( e^{t\|V(X)\|_{\mathcal{F}(\delta)}} \right) & \leq \mathbb{E} \left( e^{2t\|V_\epsilon(X)\|_{\mathcal{F}(\delta)}} \right) \\ & \leq 2 e^{2t^2 J_\delta(\mathcal{F})^2} \mathbb{E} \left[ e^{\frac{2t^2}{\sqrt{N}} r_\delta(\mathcal{F})^2 \|\bar{V}(X)\|_{\mathcal{G}(\delta)}} \right] \end{aligned}$$

Using (8.3), we conclude that

$$\mathbb{E} \left( e^{t\|V(X)\|_{\mathcal{F}(\delta)}} \right) \leq 4 \exp \left( 2t^2 J_\delta(\mathcal{F})^2 + \frac{(t r_\delta(\mathcal{F}))^4}{N} \pi_\psi \left( \|\bar{V}(X)\|_{\mathcal{G}(\delta)} \right)^2 \right)$$

Our next objective is to estimate the quantity  $\pi_\psi \left( \|\bar{V}(X)\|_{\mathcal{G}(\delta)} \right)$ . To this end, we let  $\{h^1, \dots, h^{n_{\epsilon/16}}\} \subset \mathcal{F}(\delta)$  be the centers of  $n_{\epsilon/16} = \mathcal{N}(\mathcal{F}(\delta), \mathbb{L}_2(m(x)), \epsilon/16)$   $\mathbb{L}_2(x)$ -balls of radius at most  $(\epsilon/16)$  covering  $\mathcal{F}(\delta)$ . Using the decomposition

$$(h_1 - h_2)^2 - (h^i - h^j)^2 = [(h_1 - h^i) + (h^j - h_2)] [(h_1 - h_2) + (h^i - h^j)]$$

we prove that

$$|(h_1 - h_2)^2 - (h^i - h^j)^2| \leq 8 [|h_1 - h^i| + |h^j - h_2|]$$

for any  $h_i, h^j \in \mathcal{F}(\delta)$  ( $\Rightarrow \|h_i\| \vee \|h^j\| \leq 2$ ). Using these estimates, we prove that

$$\mathcal{N}(\mathcal{G}(\delta), \mathbb{L}_2(m(x)), \epsilon) \leq \mathcal{N}(\mathcal{F}(\delta), \mathbb{L}_2(m(x)), \epsilon/16)^2 \leq \mathcal{N}(\mathcal{F}, \epsilon/32)^4$$

On the other hand, we have

$$\sup_{g \in \mathcal{G}(\delta)} \|g\| \leq 4 \sup_{h \in \mathcal{F}(\delta)} \|h\|^2 \leq 16$$

This implies that

$$I(\mathcal{G}(\delta)) = \int_0^{32} \sqrt{\log \mathcal{N}(\mathcal{G}(\delta), \epsilon)} d\epsilon \leq c I(\mathcal{F})$$

and by lemma 8.1, we can prove that

$$\pi_\psi(\|V(X)\|_{\mathcal{H}}) \leq c \int_0^{2H} \sqrt{\log \mathcal{N}(\mathcal{H}, \epsilon)} d\epsilon$$

for any class of functions  $\mathcal{H}$  s.t.  $\sup_{h \in \mathcal{H}} \|h\| \leq H$ . One concludes that

$$\pi_\psi(\|V(X)\|_{\mathcal{G}(\delta)}) \leq c I(\mathcal{F})$$

and therefore

$$\begin{aligned} \pi_\psi(\|\bar{V}(X)\|_{\mathcal{G}(\delta)}) &\leq a(\mathcal{F}) := c I(\mathcal{F}) + \tau(c I(\mathcal{F})) \\ \mathbb{E} \left( e^{t\|V(X)\|_{\mathcal{F}(\delta)}} \right) &\leq 4 \exp \left( 2t^2 J_\delta(\mathcal{F})^2 + \frac{(t r_\delta(\mathcal{F}))^4}{N} a(\mathcal{F})^2 \right) \end{aligned}$$

The end of the proof of the Laplace estimates is now easily completed. This ends the proof of the theorem. ■

## References

- [1] The large deviation principle for stochastic processes I and II. *Theory of Probability and its Applications*. 47, 567-583 and 48, 19-44 (2003).
- [2] M.A. Arcones. Moderate deviations of empirical processes. *Stochastic inequalities and applications*. Progr. Probab. 56, Birkhauser, Basel, 189-212 (2003).
- [3] A. de Acosta. Moderate Deviations for Empirical Measures of Markov Chains: Lower Bounds. *The Annals of Probability*, Vol. 25, No. 1, pp. 259-284 (1997).
- [4] A. de Acosta and X. Chen. Moderate Deviations for Empirical Measures of Markov Chains: Upper Bounds. *Journal of Theoretical Probability*, Vol. 11, No. 4 (1998)
- [5] Del Moral, P., *Feynman-Kac formulae. Genealogical and interacting particle systems with applications*, Probability and its Applications, Springer Verlag, New York (2004).

- [6] P. Del Moral, and M. Ledoux. On the Convergence and the Applications of Empirical Processes for Interacting Particle Systems and Nonlinear Filtering. *Journal of Theoretical Probability*, Vol. 13, No. 1, 225-257 (2000).
- [7] P. Del Moral, E. Rio, Concentration inequalities for Mean Field Particle Models. HAL-INRIA publication no. 6901 (29p.), April (2009).
- [8] Del Moral, P., Doucet, A., Jasra, A., *Sequential Monte Carlo Samplers*. Journal of the Royal Statistical Society, Series B, vol. 68, no. 3, pp. 411-436 (2006).
- [9] R. Douc, A. Guillin and J. Najim. Moderate Deviations for Particle Filtering. *The Annals of Applied Probability*, Vol. 15, No. 1B, pp. 587-614 (2005).
- [10] Doucet A., de Freitas J.F., Gordon N.J., *Sequential Monte-Carlo Methods in Practice*, Springer Verlag New York (2001).
- [11] F.Q. Gao. Moderate deviations for martingales and mixing random processes. *Stochastic Process. Appl.*, 61, 263-275 (1996).
- [12] F.Q. Gao. Moderate deviations and large deviations for kernel density estimators. *J. Theoret. Probab.*, 16, 401-418 (2003).
- [13] H. Djellout and A. Guillin. Moderate deviations for Markov chains with atom *Stochastic processes and their applications*. vol. 95, no2, pp. 203-21 (2001)
- [14] M. Ledoux, M. Talagrand. *Probability in Banach Spaces, Isometry and Processes*. Springer-Verlag Berlin (1991).
- [15] M. Ledoux. Sur les déviations modérées de sommes de variable aléatoires vectorielles indépendantes de même loi. *Ann. Inst. Henri Poincaré*. 28, 267-280 (1992).
- [16] D. Revuz. *Markov Chains*, North-Holland , 1976.
- [17] M. Talagrand. Sharper bounds for Gaussian and empirical processes, *Ann. Probab.* 22 (1994) 28-76.
- [18] L. M. Wu. Large deviations, moderate deviations and LIL for empirical processes. *The Annals of Probability*, vol. 22, no. 1, pp.17-27 (1994).
- [19] L. M. Wu, Forward-Backward martingale decomposition and compactness results for additive functionals of stationary ergodic Markov processes. *Annales de l'I.H.P., section B, tome 35, no 2*, pp. 121-141 (1999).
- [20] A. W. van der Vaart, J. A. Wellner. *Weak convergence of stochastic processes*. Springer Series in Statistics, Springer (1996).