Statistica Sinica **25** (2015), 921-951
doi:http://dx.doi.org/10.5705/ss.2013.134

MODERATE DEVIATIONS FOR INTERACTING PROCESSES

Pierre Del Moral, Shulan Hu and Liming Wu

University of New South Wales, Zhongnan University of Economics and Law and Université Blaise Pascal

Abstract: This article is concerned with moderate deviation principles of a class of interacting empirical processes. We derive an explicit description of the rate function, and we illustrate these results with Feynman-Kac particle models arising in nonlinear filtering, statistical machine learning, rare event analysis, and computational physics. We discuss functional moderate deviations of the occupation measures for both the strong τ-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, yielding the first results of this type for mean field interacting processes. Our approach is based on an original semigroup analysis combined with Orlicz norm inequalities, stochastic perturbation techniques, and projective limit large deviation methods.

Key words and phrases: Convergence of empirical processes, exponential inequalities, Feynman-Kac particle models, functional central limit theorems, interacting empirical processes, large deviations for projective limits, moderate deviations, sequential Monte Carlo methods.

1. Introduction

1.1. Sequential Monte Carlo methodologies

Suppose we are given a complex target probability measure π_T defined on some product space $E = S^{T+1}$, for some dimension parameter $T \geq 0$, and some measurable state space S. Stochastic particle methodologies, also termed Sequential Monte Carlo samplers (*abbreviated SMC*), consist of sampling approximately from a sequence of "interpolating" probability distributions π_n on the state spaces $E_n := S^{n+1}$ with increasing dimension, $0 \leq n \leq T$, starting from some probability measure π_0 up to the desired target measure π_T. In the further development of this section, $\boldsymbol{dx_n}$ stands for an infinitesimal neighborhood of the path sequence $\boldsymbol{x_n} = (x_0, \ldots, x_n)$, with $0 \leq n \leq T$.

We assume that these bridging measures are connected by the formula

$$\forall 0 \leq n \leq T, \qquad \pi_{n+1}(\boldsymbol{dx_{n+1}}) \;\propto\; P_{n+1}(\boldsymbol{x_n}, dx_{n+1}) \times \pi_n(\boldsymbol{dx_n}) \tag{1.1}$$

for some conditional probability distributions $P_{n+1}(\boldsymbol{x_n}, dx_{n+1})$. Here, dx_{n+1} stands for some infinitesimal neighborhood of the state $x_{n+1} \in E_{n+1}$. We choose a sequence of dominating importance sampling distribution $M_{n+1}(\boldsymbol{x_n}, dx_{n+1})$ s.t.

$$P_{n+1}(\boldsymbol{x_n}, dx_{n+1}) \ll M_{n+1}(\boldsymbol{x_n}, dx_{n+1})$$

and we set

$$\eta_{n+1}(\boldsymbol{dx_{n+1}}) := M_{n+1}(\boldsymbol{x_n}, dx_{n+1}) \times \pi_n(\boldsymbol{dx_n}). \tag{1.2}$$

We consider the corresponding importance weight function G_{n+1} defined by

$$G_{n+1}(\boldsymbol{x_{n+1}}) \propto \frac{d\pi_{n+1}}{d\eta_{n+1}}(\boldsymbol{x_{n+1}}) = \frac{dP_{n+1}(\boldsymbol{x_n}, \cdot)}{dM_{n+1}(\boldsymbol{x_n}, \cdot)}. \tag{1.3}$$

We also consider a dominating importance sampling distribution η_0 s.t. $\pi_0 \ll \eta_0$, and we set $G_0 \propto d\pi_0/d\eta_0$, the corresponding Radon-Nydodim weight function. To avoid unnecessary technical discussions, we further assume that the functions G_n are positive and bounded. For more general models, including indicator type functions and unbounded functions G_n, we refer the reader to Section 7.2.1 in the research monograph of Del Moral (2004).

The SMC algorithm is a population type Monte Carlo algorithm based on sampling sequentially a collection of N random trajectories. Initially, we sample $(X_0^i)_{1\leq i\leq N}$ with some proposal distribution $\eta_0(dx_0)$. Then we resample $(\widehat{X}_0^i)_{1\leq i\leq N}$ with the discrete measure $\propto \sum_{1\leq i\leq N} G_0(X_0^i)\delta_{X_0^i}$ on E_0. For each of the selected variables $\widehat{X}_0^i$, we sample a random variable $\tilde{X}_1^i$ with the proposal distribution $M_1(\widehat{X}_0^i, dx_1)$, and we set $X_1^i = \left(\widehat{X}_0^i, \tilde{X}_1^i\right)$. Then we resample $(\widehat{X}_1^i)_{1\leq i\leq N}$ with the discrete measure $\propto \sum_{1\leq i\leq N} G_1(X_1^i)\delta_{X_1^i}$ on the product space E_1. For each of the selected variables $\widehat{X}_1^i$, we sample a random variable $\tilde{X}_2^i$ with the proposal distribution $M_2(\widehat{X}_1^i, dx_2)$, and we set $X_2^i = \left(\widehat{X}_1^i, \tilde{X}_2^i\right)$, and so on.

By (1.3) we have $P_{n+1}(\boldsymbol{x_n}, dx_{n+1}) \propto\ M_{n+1}(\boldsymbol{x_n}, dx_{n+1})\ G_{n+1}(\boldsymbol{x_{n+1}})$. Thus, combining (1.1) and (1.2), we readily check that

$$\pi_n(\boldsymbol{dx_n}) \propto G_n(\boldsymbol{x_n})\ \eta_n(\boldsymbol{dx_n}) \text{ and } \eta_{n+1}(\boldsymbol{dx_{n+1}}) := M_{n+1}(\boldsymbol{x_n}, dx_{n+1})\ \ \pi_n(\boldsymbol{dx_n}).$$

This shows that the interpolating sequence of distributions satisfies the nonlinear equation

$$\eta_n \longrightarrow (\pi_n =)\ \ \Psi_{G_n}(\eta_n) \longrightarrow \eta_{n+1} := \Phi_{n+1}(\eta_n) := \Psi_{G_n}(\eta_n)\boldsymbol{M_{n+1}} \tag{1.4}$$

defined in terms of two operator:

- The Boltzmann-Gibbs transformation Ψ_{G_n} from the set of probability measures on E_n into itself, defined by

$$\eta_n(\boldsymbol{dx_n}) \rightsquigarrow \Psi_{G_n}(\eta_n)(\boldsymbol{dx_n}) := \frac{G_n(\boldsymbol{x_n})}{\eta_n(G_n)}\ \eta_n(\boldsymbol{dx_n})$$

and

$$\eta_n(G_n) := \int \eta_n(\boldsymbol{dx_n})\ G_n(\boldsymbol{x_n}).$$

- The Markov transport equation from the set of probability measures on E_n into the set of probability measures on E_{n+1}, defined by

$$\pi_n(\boldsymbol{dx_n}) \rightsquigarrow (\pi_n \boldsymbol{M_{n+1}})(\boldsymbol{dx_{n+1}}) := \int \pi_n(\boldsymbol{dx_n})\boldsymbol{M_{n+1}}(\boldsymbol{x_n}, \boldsymbol{dx_{n+1}}).$$

In this framework, with some obvious abusive notation the SMC algorithm discussed above takes the following synthetic form

$$\left(X_n^1, \dots, X_n^N\right) \text{ i.i.d. } \sim \Phi_n\Big(\frac{1}{N}\sum_{1\le i\le N} \delta_{X_{n-1}^i}\Big). \tag{1.5}$$

The rationale behind these interacting processes is that if the empirical measure $\eta_{n-1}^N := \frac{1}{N}\sum_{1\le i\le N}\delta_{X_{n-1}^i} \simeq \eta_{n-1}$ for N large enough at a given rank $(n-1)$, then at rank n we have $\left(X_n^1, \dots, X_n^N\right)$ are almost i.i.d. samples of the distribution η_n. From a statistical point of view, the interacting particle evolution equation (1.5) can be also be interpreted as a kinetic transformed statistical model defining a sequence of interacting empirical measures η_n^N. At any time n, the desired measures η_n are approximated by the empirical measures η_n^N. More general particle models are discussed in Section 1.2, dedicated to mean field particle approximations of a general class of measure-valued processes.

The measure-valued equations and the corresponding particle algorithm arise in such applications areas, as physics, biology, and advanced stochastic engineering sciences. For instance, in signal processing, the conditional distributions of the paths of the Markov signal $\boldsymbol{X_n} := (X_0, \dots, X_n)$, given a series of noisy observations $(Y_0, \dots, Y_n) = (y_0, \dots, y_n)$, satisfy a two-step prediction-updating equation of the form (1.6): for any $n \ge 0$,

$$\eta_n = \text{Law}\left(\boldsymbol{X_n} \mid \forall 0 \le p < n \ \ Y_p = y_p\right).$$

In this context, the $G_n(\mathbf{x_n}) = p_n(y_n \mid \mathbf{x_n},\ y_0, \dots, y_{n-1})$ are given by the likelihood functions of the observation y_n w.r.t. the signal sequence $\mathbf{x_n}$, and the Boltzmann-Gibbs transformation coincides with the Bayes' rule. The updating-prediction evolution equation (1.6) is sometimes called the discrete generation nonlinear filtering equation; Here the particle model is also called a particle filter or a genetic algorithm.

In the context of sequential bayesian inference, the distributions $\eta_n(d\theta)$ could also be the posterior distributions of an unknown parameter given the data collected up to time n,

$$\eta_n(d\theta) = p(\theta \mid y_p,\ p < n) \propto \left\{ \prod_{0 \le k < n} p(y_k \mid \theta, y_0, \ldots, y_{k-1}) \right\} p(d\theta).$$

For any Markov chain Monte Carlo (*abbreviated MCMC*) transition M_n with target measure $\eta_n = \eta_n M_n$ we clearly have that

$$G_{n-1}(\theta) = p(y_n \mid \theta,\ y_0, \ldots, y_{n-1}) \Rightarrow \eta_n = \Phi_n\left(\eta_{n-1}\right) := \Psi_{G_{n-1}}(\eta_{n-1})M_n.$$

Applying the particle methodology (1.5) we design an interacting type particle MCMC model that approximate sequentially the desired posterior distribution $p(\theta \mid y_0, \ldots, y_n)$ by an interacting empirical process, as $N \uparrow \infty$. When the functions $G_n(\theta)$ are unknown we consider the extended model $\overline{\theta} = (\theta, X_k^{\theta,i}, 1 \le i \le N',$ $0 \le k \le T)$, where the latent variables $X_k^{\theta,i}$ stands for the particle filter discussed above associated with the fixed parameter θ. In this context, it is more or less well known that the θ-marginal of the extended distribution

$$\begin{aligned}\overline{\eta}_n(d\overline{\theta}) \propto \left\{ \prod_{1 \le k \le n} \overline{G}_{k-1}(\overline{\theta}) \right\} p(d\overline{\theta}) \quad \text{with} \quad \overline{G}_{n-1}(\overline{\theta}) \\ = \frac{1}{N'} \sum_{1 \le i \le N'} p_n(y_n \mid \boldsymbol{X_n^{\theta,i}},\ y_0, \ldots, y_{n-1})\end{aligned}$$

coincides with the desired posterior distribution $p(\theta \mid y_0, \ldots, y_n)$, for any $N' \ge 1$. A detailed proof of this result can be found in Del Moral (2013). In this framework, for any Markov chain Monte Carlo (*abbreviated MCMC*) transition M_n with target measure $\overline{\eta}_n = \overline{\eta}_n \overline{M}_n$, we clearly have that

$$\overline{\eta}_n = \Phi_n\left(\overline{\eta}_{n-1}\right) := \Psi_{\overline{G}_{n-1}}(\overline{\eta}_{n-1})\overline{M}_n.$$

Applying the particle methodology (1.5) we design an interacting MCMC sampler coupled with particle filters that approximate sequentially the desired posterior distribution $p(\theta \mid y_0, \ldots, y_n)$ by an interacting empirical process, as $N \uparrow \infty$.

1.2. Mean field interacting particle processes

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

$$\eta_{n+1} = \Phi_{n+1}(\eta_n). \tag{1.6}$$

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

$$\Phi_n(\eta_{n-1}) = \eta_{n-1} K_{n,\eta_{n-1}} \tag{1.7}$$

for some collection of Markov kernels $K_{n,\mu_{n-1}}$ indexed by the time parameter n and the set of probability measures μ_{n-1} on the space E_{n-1}. These models provide a natural interpretation of the distribution laws η_n as the laws of a non-linear Markov chain whose elementary transitions depend on the current distribution. In further developments, we assume that the mappings $\left(x_n^i\right)_{1\leq i\leq N} \in E_n^N \mapsto K_{n+1,\frac{1}{N}\sum_{j=1}^N \delta_{x_n^j}}\left(x_n^i, A_{n+1}\right)$ 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 $\xi_n^{(N)} = \left(\xi_n^{(N,i)}\right)_{1\leq i\leq N}$, with elementary transitions defined as

$$\mathbb{P}\left(\xi_{n+1}^{(N)} \in dx \;\middle|\; \mathcal{A}_n^{(N)}\right) = \prod_{i=1}^N K_{n+1,\eta_n^N}(\xi_n^{(N,i)}, dx^i) \quad \text{with} \quad \eta_n^N := \frac{1}{N}\sum_{j=1}^N \delta_{\xi_n^{(N,j)}}. \tag{1.8}$$

Here, $\mathcal{A}_n^{(N)} := \sigma\left(\xi_p^{(N)},\ 0\leq p\leq n\right)$ stands for the sigma-field generated by the random variables $(\xi_p^{(N)})_{0\leq p\leq n}$, and $dx = dx^1\times\ldots\times dx^N$ stands for an infinitesimal neighborhood of a point $x=(x^1,\ldots,x^N)\in E_n^N$. The initial system $\xi_0^{(N)}$ consists of N independent and identically distributed random variables with common law η_0. To simplify the presentation, when there is no possible confusion we suppress the parameter N, so that we write ξ_n and ξ_n^i instead of $\xi_n^{(N)}$ and $\xi_n^{(N,i)}$. For a thorough description of these discrete generation and nonlinear McKean type models, we refer the reader to Del Moral (2004).

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 and statistical machine learning with the corresponding particle filter technology, in stochastic engineering with interacting type Metropolis and Gibbs sampler methods, and in statistical physics with quantum and diffusion Monte Carlo algorithms leading to precise estimates of the top eigenvalues and the ground states of Schrödinger operators in Hetherington (1984), Caffarel (1989), Assaraf and Caffarel (2000), Assaraf, Caffarel, and Khelif (2000), Caffarel et al. (2006) and Caffarel (2011). For a more detailed discussion on these application areas, we again refer the reader to Doucet, de Freitas, and Gordon (2001), Del Moral (2004, 2013), Del Moral, Doucet, and Jasra (2006), and the references therein.

A typical example 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

$$\Phi_{n+1}(\eta_n)(dy) = (\Psi_{G_n}(\eta_n) M_{n+1})(dy) := \int \Psi_{G_n}(\eta_n)(dx)\ M_{n+1}(x,dy). \quad (1.9)$$

In this situation, the flow of measures η_n is given for any bounded measurable function f on E_n as

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

where X_n stands for a Markov chain with initial distribution η_0 and Markov transitions M_n.

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

$$\Psi_{G_n}(\eta_n)=\eta_n S_{\eta_n,G_n} \text{ with } S_{\eta_n,G_n}(x,dy)=G_n(x)\delta_x(dy)+(1-G_n(x))\,\Psi_{G_n}(\eta_n)(dy), \quad (1.10)$$

then we find that $K_{n+1,\eta_n} = S_{\eta_n,G_n} M_{n+1}$.

The mathematical and numerical analysis of these mean field particle models (1.8) is one of the most active research subject in pure and applied probability, as well as in statistical machine learning, 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.

Moderate deviation properties can be thought as an intermediate asymptotic estimation between the central limit theorem and the large deviations principles. The theory of moderate and large deviations is a wide and fast developing branch of probability and statistic theory. Nevertheless, to the best of our knowledge the existing literature on moderate deviation principles is concerned with independent and identically distributed random sequences, Markov chain processes, and random fields models; see for instance the article by Dobrushin and Shlosman (1994), the series of works by Ledoux (1992), Wu (1994, 1999), Gao (1996, 2003), de Acosta (1997), de Acosta and Chen (1998), Arcones (2003a,b), Gao and Zhao (2011), and the more recent and seminal works of Wu and Zhao (2008), and Peligrad et al. (2013).

Surprisingly, moderate deviations for mean field type interacting particle systems have not been covered by the literature. In the present article, we analyze these questions using a stochastic perturbation analysis of an abstract class of nonlinear semigroups in distribution spaces. The central idea is to use backward semigroup expansions to express any global error quantity in terms of the local sampling errors induced by the mean field simulation. We use the differential calculus on measure spaces developed in the recent article of the first author with E. Rio (2011). These authors generalize the classical Hoeffding, Bernstein and these Bennett inequalities for independent random sequences to interacting particle systems, but leave open the question of moderate deviation principles. In the context of continuous time mean field particle models, the analysis of the first and the second order variational derivatives of limiting nonlinear semigroups w.r.t. the initial data has also been developed by Kolokoltsov (2007, 2010, 2013) to analyze the smoothness of nonlinear semigroups, and to derive dynamical law of large numbers and fluctuation theorems.

We complete this study with functional moderate deviations of mean field particle models for both the τ-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. The mathematical framework developed in Del Moral and Rio (2011), and in the present work applies to a general class of mean field particle models, including Feynman-Kac integration models, interacting jump processes, McKean Vlasov diffusion type models, as McKean collision type models of gases. For a detailed derivation of these application models, we refer to Del Moral and Rio (2011), Del Moral (2013).

1.3. Outline of the paper

The main results are presented in Section 2 : the moderate deviation principles (MDP) in finite dimension ; in infinite dimension but in the τ-topology; for empirical process indexed by a class of functions; We describe some main arguments 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. The proofs of some technical results are provided in a Web-Appendix.

1.4. Some notation

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 ν with $\nu(E)=0$, and the Banach space of all bounded and measurable functions f equipped with the uniform norm $\|f\|$. We denote by

$\mathrm{Osc}_1(E)$ the convex set of $\mathcal{E}$-measurable functions f with oscillations $\mathrm{osc}(f) := \sup_{x\neq y}|f(x)-f(y)| \leq 1$. 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 $\mu \in \mathcal{M}(E)$. 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_1M_2) denote the composition operator $(M_1M_2)(f) = M_1(M_2(f))$. For time homogenous state spaces, we denote by $M^m = M^{m-1}M = MM^{m-1}$ 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 $(F,\mathcal{F})$ 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 on

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

For the bounded integral operator M with $M(1)(x) = M(1)(y)$ 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

$$\beta(M) := \sup\ \{\mathrm{osc}(M(f))\ ;\ \ f \in \mathrm{Osc}_1(F)\}. \tag{1.11}$$

Finally, we let $\Phi_{p,n}$, $0 \leq p \leq n$, be the semigroup associated with the measure valued equation defined in (1.6), $\Phi_{p,n} = \Phi_n \circ \Phi_{n-1} \circ \cdots \circ \Phi_{p+1}$. For $p = n$, we use the convention $\Phi_{n,n} = Id$, the identity operator.

2. 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

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

where

(i) the first order operators $(\mathcal{D}_\eta\Phi)_{\eta\in\mathcal{P}(E_1)}$ 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

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

(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

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

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

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

We observe that any mapping $\Phi\in\Upsilon(E_1,E_2)$ is Fréchet differentiable in the sense that

$$\Phi(\mu)-\Phi(\eta)-(\mu-\eta)D_{\eta}\Phi = \mathrm{o}(\mu-\eta),$$

where $\lim_{\mu\to\eta}\|(\mu-\eta)\|_{\mathrm{tv}}^{-1}\|\mathrm{o}(\mu-\eta)\|_{\mathrm{tv}}$. This implies that Φ is also Gâteaux differentiable at any $\mu\in\mathcal{P}(E_1)$ in any direction $\nu=(\eta-\mu)\in\mathcal{M}_0(E_1)$, with $\eta\in\mathcal{P}(E_1)$, in the sense that

$$\lim_{\epsilon\downarrow 0}\left\|\frac{1}{\epsilon}\left[\Phi(\mu+\epsilon\nu)-\Phi(\mu)\right]-\nu D_{\mu}\Phi\right\|_{\mathrm{tv}} = 0.$$

One practical way to compute the integral operator $D_{\mu}\Phi$ is to check that for any $f\in\mathcal{B}_b(E_2)$ we have that

$$\frac{d}{d\epsilon}\Phi(\mu+\epsilon\nu)(f)_{|\epsilon=0} = \nu D_{\mu}\Phi(f).$$

Inversely, let Φ be a Gâteaux differentiable mapping at any $\mu\in\mathcal{P}(E_1)$ in any direction $\nu=(\eta-\mu)\in\mathcal{M}_0(E_1)$, with $\eta\in\mathcal{P}(E_1)$, such that

$$\lim_{\epsilon\downarrow 0}\left\|\frac{1}{\epsilon}\left[\nu D_{\mu+\epsilon\nu}\Phi-\nu D_{\mu}\Phi\right]-\nu^{\otimes 2}D_{\mu}^2\Phi\right\|_{\mathrm{tv}} = 0$$

for some bounded integral operator $D_{\mu}^2\Phi$ from $\mathcal{B}_b(E_2)$ into $\mathcal{B}_b(E_1\times E_1)$. Here again, one way to compute $D_{\mu}^2\Phi$ is to check that

$$\frac{d^2}{d\epsilon^2}\Phi(\mu+\epsilon\nu)(f)_{|\epsilon=0} = \frac{d}{d\epsilon}\nu D_{\mu+\epsilon\nu}\Phi(f)_{|\epsilon=0} = \nu^{\otimes 2}D_{\mu}^2\Phi(f).$$

We further assume that the mappings $(\nu,\mu)\mapsto\nu^{\otimes 2}D_{\mu}^2\Phi$ and $(\nu,\mu)\mapsto\nu D_{\mu}\Phi$ are continuous. In this situation, Φ is C^2-Gâteaux differentiable mapping at any $\mu\in\mathcal{P}(E_1))$ in any direction $\nu=(\eta-\mu)\in\mathcal{M}_0(E_1)$, with $\eta\in\mathcal{P}(E_1)$, and we have Taylor's theorem with integral remainder

$$\Phi(\mu+\nu) = \Phi(\mu)+\nu D_{\mu}\Phi+\int_0^1 \; (1-t)\;\nu^{\otimes 2}D_{\mu+t\nu}^2\Phi \;\; dt.$$

The r.h.s. integral is the Gelfand-Pettis weak sense integral in Brooks (1969). This yields the first order decomposition

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

with the second order remainder measure

$$\mathcal{R}^\Phi(\eta, \mu) = \int_0^1 \ (1-t)\ (\eta - \mu)^{\otimes 2} D^2_{\mu + t(\eta - \mu)} \Phi \ \ dt.$$

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

$$(K) \qquad \| \left[K_\mu - K_\eta \right](f) \| \leq \int \ |(\mu - \eta)(h)| \ T^K_\eta(f, dh). \tag{2.5}$$

Here T^K_η stands for some collection of bounded integral operators from $\mathcal{B}(E_2)$ into $\mathcal{B}(E_1)$ such that

$$\sup_{\eta \in \mathcal{P}(E_1)} \int \ \mathrm{osc}(h)\ T^K_\eta(f, dh) \leq \ \mathrm{osc}(f)\ \delta\left(T^K\right) \tag{2.6}$$

for some finite constant $\delta\left(T^K\right) < \infty$. With $K_\eta(x, dy) = \Phi(\eta)(dy)$, 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 Φ. In this situation, we denote by (Φ) the corresponding condition; and whenever it is met, we say that the mapping Φ satisfy condition (Φ).

Throughout this paper we assume

(H1) The given collection of McKean transitions $K_{n,\eta}$ satisfies (2.5) and (2.6), and 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 (1.6) 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, assumptions can be found in Del Moral and Rio (2011) and Del Moral (2013), 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

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

with $G_{n,\eta} = G_n/\eta(G_n)$ and $M_{n+1,\eta}(f) := M_{n+1}(f) - \Phi_{n+1}(\eta)(f)$. Assuming that $g_n = \sup_{x,y} G_n(x)/G_n(y) < \infty$, we find the Lipschitz estimates

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

as well as the first order estimation

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

with the first order functional $D_\eta \Phi_{n+1}(f) = G_{n,\eta} \ M_{n+1,\eta}(f)$.

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. 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 below.

Definition 1. (V_n^N, W_n^N) is the sequence of random fields given by the pair of stochastic perturbation formulae

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

where $\eta_n^N = (1/N) \sum_{j=1}^N \delta_{\xi_n^{(N,j)}}$ is the empirical distribution of ξ_n^N.

The sequence V_n^N is sometimes referred to the local sampling random field model. The centered random fields V_n^N have conditional variance functions given by

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

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

$$\Phi_n(\eta) - \Phi_n(\mu) \simeq (\eta - \mu) D_\mu \Phi_n \tag{2.10}$$

with a first order integral operator $D_\mu \Phi_n$ from $\mathcal{B}(E_n)$ into $\mathcal{B}(E_{n-1})$.

Definition 2. $(\mathcal{D}_{p,n})_{0 \leq p \leq n}$ is 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

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

we proved in Del Moral and Rio (2011) 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, centered Gaussian 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 $n \geq 0$,

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

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

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

Here we analyze asymptotic expansions for probabilities of moderate deviations. We first give the description of a large deviation principle (LDP).

Definition 3. With $(\alpha(N))_{N\geq 1}$ a sequence of positive numbers such that $\lim_{N\to\infty} \alpha(N) = \infty$, 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 good rate function $I : x \in S \mapsto I(x) \in [0, \infty]$ if: for every finite constant $a < \infty$, the level sets $\{x \in S : I(x) \leq a\}$ are compact sets; for each $A \in \mathcal{S}$,

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

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

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

Theorem 1. *For any nondecreasing function $\alpha(N)$ such that $\lim_{N\to\infty} \alpha(N)/N = 0$, any $n \geq 0$, and any collection of functions $f_p \in \mathcal{B}(E_p)$, $0 \leq p \leq n$, we have*

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

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

The proof is in Section 3.2 and Section 3.3. Here (2.15) is a more or less direct consequence of (2.14) combined with(2.11).

Now, for any finite subset $\mathcal{F}_n = \{f_n^1, \ldots, f_n^{d_n}\} \subset \mathcal{B}(E_n)^{d_n}$, with $d_n \geq 1$, consider the projection mapping

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

By a theorem of J. Gärtner and R.S. Ellis, using (2.15), we prove the following.

Corollary 1. *The random sequence $\pi_{\mathcal{F}_n}(W_n^N)$ satisfies an MDP principle in $\mathbb{R}^{d_n}$ with speed $\alpha(N)$ and good rate function given for any $v \in \mathbb{R}^{d_n}$, since*

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

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

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

Using (2.14), we readily prove the following corollary.

Corollary 2. *The random sequences $\left[\pi_{\mathcal{F}_0}\left(V_0^N\right), \ldots, \pi_{\mathcal{F}_n}\left(V_n^N\right)\right]$ satisfy a MDP principle in $\mathbb{R}^{d_0+\ldots+d_n}$ with speed $\alpha(N)$, and the good rate function given for any $v = (v_0, \ldots, v_n) \in \mathbb{R}^{d_0+\ldots+d_n}$ since*

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

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.

We strengthen these MDP in two ways, firstly, deriving the MDP for the random fields sequences on the set of measures equipped with the τ topology.

Theorem 2. *When the state spaces E_n are Polish spaces (metric, complete and separable), the sequence of random fields $\left(V_0^N, \ldots, V_n^N\right)$ satisfies an MDP in the product space $\prod_{p=0}^n \mathcal{M}(E_p)$ equipped with the product τ topology, with speed $\alpha(N)$ and 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*

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

where the good rate functions I_n on $\mathcal{M}(E_n)$ are for any $\mu_n \in \mathcal{M}(E_n)$,

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

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

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

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.

Our second main result is a functional moderate deviation for stochastic processes indexed by a separable collection $\mathcal{F}_n$ of measurable functions $f_n : E_n \to \mathbb{R}$ such that $\|f_n\| \leq 1$. We let $l_\infty(\mathcal{F}_n)$ be the space of all bounded real functions $F_n : f_n \in \mathcal{F}_n \mapsto F_n(f_n) \in \mathbb{R}$ on $\mathcal{F}_n$ with the supnorm $\|F_n\|_{\mathcal{F}_n} = \sup_{f_n \in \mathcal{F}_n} |F_n(f_n)|$. 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 $\mathcal{F}_n$, consider 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 $I(\mathcal{F}_n)$ we denote the uniform covering numbers and entropy integral given by

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

We assume the following

(A1) $\mathcal{N}(\epsilon, \mathcal{F}_n) < \infty$ for any $\epsilon > 0$, 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 η on E_n. Various classes of functions with finite covering and entropy integral are given in van der Vaart and Wellner (1996).

For any $\delta > 0$, we 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 a separable collection $\mathcal{F}_n'$ of measurable functions f_n on E_n, $\|f_n\| \leq 1$, such that $I(\mathcal{F}_n') < c_0(n)\ I(\mathcal{F}_{n+1})$ and, for any probability measure μ, any $\delta > 0$,

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

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

In the context of the Feynman-Kac models, using (2.7), we find that

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

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

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

Using elementary manipulations, we show that (**A2**) is met with $c_1(n)=1/(2\sqrt{g_n}) \le 1$, $c_2(n) = 2g_n^2$, 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 Del Moral and Ledoux (2000), 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 $\mathcal{G}_n \subset \mathcal{F}_n$, 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 $\pi_{\mathcal{F}_n,\mathcal{G}_n}(\nu)(g) = v(g)$, 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 3. *If the class of observables $\mathcal{F}_n$ satisfies (A1) and (A2), the sequence of stochastic processes W_n^N satisfies the large deviation principle in $\mathcal{L}_\infty(\mathcal{F}_n)$ with the good rate function $I_{\mathcal{F}_n}^{W_n}$,*

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

where J_n is given in (2.18).

For finite sets $\mathcal{F}_n$, this reduces to the MDP presented in (2.16). The τ-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 τ-topology. Then Theorem 3 is a direct consequence of Theorem 2. In more general situations, by Wu (1994) or a theorem of Arcones (2003b), 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 exponential asymptotic equicontinuity condition

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

with the collection of functions

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

The proof is given in Section 5.

3. Asymptotic Laplace Expansions

3.1. Some preliminary results

Lemma 1. *For any $0 \le p \le n$, we have $\Phi_{p,n} \in \Upsilon(E_p, E_n)$ with the first order decomposition*

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

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 $\mathcal{R}^{\Phi_{p,n}}(\eta,\mu)$. For any $N \ge 1$, we have the first order decomposition

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

and the semigroup $(\mathcal{D}_{p,n})_{0\le p\le n}$ at (2.11).

Lemma 2. *For every $f \in \mathrm{Osc}_1(E_n)$, $N \ge 1$ and any $n \ge 0$ and $m \ge 1$, we have the $\mathbb{L}_m$ estimates*

$$\mathbb{E}\left(\left|V_n^N(f_n)\right|^m \Big|\xi_{n-1}^{(N)}\right)^{1/m} \le b(m)$$

and

$$\sqrt{N}\mathbb{E}\left(\left|\left[\eta_n^N - \eta_n\right](f_n)\right|^m\right)^{1/m} \le b(m)\ \sum_{p=0}^{n}\delta(T^{\Phi_{p,n}}), \tag{3.3}$$

as well as the bias estimate

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

A detailed proof of (3.1) is in Del Moral and Rio (2011). Formula (3.2) is a direct consequence of the inductive decomposition $W_n^N = V_n^N + W_{n-1}^N \mathcal{D}_n + \sqrt{N}\ R^{\Phi_n}\left(\eta_{n-1}^N, \eta_{n-1}\right)$.

The first estimates in Lemma 2 are direct consequence of Kintchine's inequality, and the sencond are more or less direct consequences of the Lipschitz properties of the semigroups $\Phi_{p,n}$. A proof is in Appendix A of the Web Appendix.

3.2. Second order remainder measures

This section is mainly concerned with the non asymptotic Laplace estimates of the second order remainder measures

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

Proposition 1. *For every $f \in \mathrm{Osc}_1(E_n)$, $N \geq 1$, $n \geq 0$, we have*

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

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

Proof. By (3.2),

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

Combining (3.3) with the generalized Minkowski inequality,

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

Then, for a Gaussian centered random variable with $\mathbb{E}(X^2) = 1$, $b(2m)^{2m} = \mathbb{E}(X^{2m})$ and, for any $t \in [0, 1/2[$,

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

Corollary 3. For every $f \in \mathrm{Osc}_1(E_n)$, $N \geq 1$, $n \geq 0$, and for every $\epsilon > 0$, we have

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

where

$$\delta_n(\epsilon, N) = \frac{r(n)}{\epsilon\sqrt{N}} \log\Big(1 + \frac{\epsilon\sqrt{N}}{r(n)}\Big).$$

For any nondecreasing function $\alpha(N)$ such that $\lim_{N\to\infty} \alpha(N)/N = 0$, we have

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

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

Proof. Using the fact that

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

we readily find that

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

Choosing $t = (1/2)(1 - 1/(1+\epsilon))$, we find that

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

which ends the proof.

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

Lemma 3. *Let (X_N), (Y_N) be two sequences of random valuables such that, for any $\lambda \geq 0$,*

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

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

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

3.3. Asymptotic Laplace transform estimates

This section is mainly concerned with the proof of Theorem 1. Fluctuation properties of the first order random field sequence $\sum_{p=0}^n V_p^N \mathcal{D}_{p,n}$ are encoded in a pair of martingale sequences .

We associate with collection of functions $f = (f_n)_{n\geq 0} \in \prod_{n\geq 0} \mathcal{B}(E_n)$, the pair of $\sigma\left(\xi_0^{(N)}, \ldots, \xi_n^{(N)}\right)$-martingale sequences

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

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 $N \geq 1$, the angle bracket of $M_n^{(N)}(f)$ is

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

with the random increments

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

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

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

The main object here is to prove that

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

Using the exponential martingale decomposition

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

We can prove the following estimates. A detailed proof is in Appendix A of the Web Appendix.

Lemma 4. *There exist a pair of functions* $(\tau_{j,n}^{(N)}(f))_{j=1,2}$ *that converge to* 0 *as* N *tends to* ∞*, such that*

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

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

Proposition 2. *We have*

$$\lim_{N\to\infty}\frac{1}{\alpha(N)}\log\mathbb{E}\Big(e^{\alpha(N)\big((1/\sqrt{\alpha(N)})\ M_n^{(N)}(f)-(1/2)\ \langle M^{(N)}(f)\rangle_n\big)}\Big)=0,$$

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

where $\overline{c}_n$ *stands for some finite constant* $\overline{c}_n:=\sum_{p=0}^n c(p)$ *with*

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

Before getting to the proof of the proposition, we make a couple of comments. First, replacing in (3.8) the parameter t by $(\alpha(N)/\sqrt{N})t$ we find that

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

from which we conclude that

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

The stochastic processes

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

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

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

for any $\epsilon\in[0,1]$. Here ϵf stands for the sequence of functions $(\epsilon f_p)_{0\leq p\leq n}$. Therefore (3.7) is a direct consequence of Lemma 3.

Proof of proposition 2. Since $\langle M^{(N)}(f)\rangle_n\leq\sigma_n^2(f):=\sum_{p=0}^n\operatorname{osc}(f_p)^2$, using Lemma 4, shows that

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

This is the first assertion. To prove (3.8), for every $n \geq 1$, $\eta \in \mathcal{P}(E_{n-1})$, and $f_n \in \mathcal{B}(E_n)$, set

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

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

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

This implies that

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

and therefore

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

Using (3.3), we have the upper bound

$$\sqrt{N}\,\mathbb{E}\left(|\Sigma_n(\eta_{n-1}^N, f_n) - \Sigma_n(\eta_{n-1}, f_n)|^m\right)^{1/m} \leq b(m)\; c(n),$$

and one concludes that

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

The $\mathbb{L}_m$-inequalities then imply that, for any $t > 0$,

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

Then (3.8) follows, and completes ends the proof.

3.4. Proof of Theorem 1

Proof of (2.14). This was done in Subsection 3.3.

Proof of (2.14) $\Longrightarrow$ (2.15). If a final time horizon n is fixed, then we have for any function $f_n \in \mathcal{B}(E_n)$,

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

Let (A_n^N, B_n^N) the pair of random fields

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

By (2.14), we have

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

and by (3.5)

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

This yields that

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

where (2.15) follows by Lemma 3.

4. Moderate Deviations in τ-topology

We prove Theorem 2. For this theorem we require that the state spaces E_n are Polish spaces. The τ-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 $A \in \mathcal{E}$.

We first provide a more explicit representation of the rate functions I_n appearing in (2.17). Let $K^\star_{n,\eta_{n-1}}$ 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

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

Using a spectral decomposition of the self-adjoint operator $K^\star_{n,\eta_{n-1}} K_{n,\eta_{n-1}}$ we get that

$$I_n(\mu) = \frac{1}{2} \sum_{m\geq 0} \eta_n \left[(h_\mu) \left(K^\star_{n,\eta_{n-1}} K_{n,\eta_{n-1}}\right)^m (h_\mu)\right], \tag{4.1}$$

if $\mu \in \mathcal{M}(E_n)$, $\mu(E_n) = 0$, $\mu \ll \eta_n$, $h_\mu = d\mu/d\eta_n \in \mathbb{L}_2(\eta_n)$, and $I_n(\mu) = +\infty$ otherwise. A proof of this is in Section 2.4 of the Web Appendix.

4.1. Proof of Theorem 2 by projective limit

Let $\mathcal{U}(E_n)$ be 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 $d \geq 1$, and let $\sigma(U_n)$ be the σ-field generated by U_n. Take

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

as the restriction of the measure μ to the sigma-field $\sigma(U_n)$. Here $\mathcal{M}(E_n, \sigma(U_n))$ can be identified with $\mathbb{R}^{U_n} \simeq \mathbb{R}^d$. Furthermore, the σ-algebra and the τ-topology induced on $\mathcal{M}(E_n, \sigma(U_n))$ by the restriction mapping π_{U_n} coincide with the natural topology and the Borel sigma-field on $\mathbb{R}^d$.

We say that a partition U_n' is finer than U_n, $U_n' \geq U_n$, if $\sigma(U_n') \supset \sigma(U_n)$. We let $\pi_{U_n',U_n} : \mu \in \mathcal{M}(E_n, \sigma(U_n')) \mapsto \pi_{U_n',U_n}(\mu) \in \mathcal{M}(E_n, \sigma(U_n))$ be the restriction of the measure μ 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

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

We take $\mathbf{M}(E_n)$ as the set of finite additive set functions from $\mathcal{E}_n$ into $\mathbb{R}_+$, equipped with the τ_1-topology of setwise convergence. Thus, a sequence $\mu_k \in \mathbf{M}(E_n)$ τ_1-converges to some $\mu \in \mathbf{M}(E_n)$ if $\lim_{k\to\infty} \mu_k(A) = \mu(A)$, for any $A \in \mathcal{E}_n$.

Let $\theta : \lim_{\mathcal{U}_n} \mathcal{M}_n \to \mathbf{M}(E_n)$ 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 τ_1 convergence, θ is an homeomorphism.

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

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

Since we have

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

we readily find that

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

from which we conclude that

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

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

Proposition 3. *The sequence of random fields V_n^N satisfies 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*

$$\bar{I}_n(\mu) = \sup_{U_n\in\mathcal{U}_n} I_{U_n}(\pi_{U_n}(\mu)). \tag{4.2}$$

The proof of the MDP for V_n^N is now a direct consequence of the next lemma.

Lemma 5. *The domain $\mathrm{Dom}(\bar{I}_n) = \left\{\mu\in\mathbf{M}(E_n) : \bar{I}_n(\mu) < \infty\right\}$ 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* (4.2) *coincides with I_n in* (2.17).

Remark 1. Since the relative topology on $\mathcal{M}(E_n)$ induced by the τ_1 topology coincides with the τ 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.

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

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

Proof of Theorem 2. 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 sequence

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

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

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

where $f_n^{u_n} = \sum_{i=1}^d u_n^i \, 1_{U_n^i}$. Hence by projective limit, $\left(V_0^N, \ldots, V_n^N\right)$ satisfies the MDP on $\mathbf{M}(E_0) \times \cdots \times \mathbf{M}(E_n)$ w.r.t. the product τ_1-topology, with the rate function

$$I_{[0,n]}(\mu_0, \cdots, \mu_n) = \sum_{p=0}^n \bar{I}_p(\mu_p).$$

The proof of Theorem 2 is completed by Lemma 5.

Now, we come to the

Proof of lemma 5. Consider a sequence of partitions $U_{n,d}$, finer and finer as d increases, such that $\sigma\left(\bigcup_{d\geq 1} U_{n,d}\right) = \mathcal{E}_n$. To prove that $\mathrm{Dom}(\bar{I}_n) \subset \mathcal{M}(E_n)$, we use

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

See for instance (2.5) in the Web Appendix. Therefore $\left\{d\pi_{U_{n,d}}(\mu)/d\pi_{U_{n,d}}(\eta_n)\right\}_{d\geq 1}$ is a $\mathbb{L}_2$-bounded martingale w.r.t. the probability measure η_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

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

in $\mathbb{L}_2(\eta_n)$, as d goes to infinity. We show that h_μ 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 a partition $V_{n,d}$ which is finer than $U_{n,d}$ and $U'_{n,d}$, and such that $V_{n,d+1}$ is finer than $V_{n,d}$. As above,

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

in $\mathbb{L}_2(\eta_n)$, as $d \to \infty$. 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 $h_\mu = \tilde{h}_\mu$, $\eta_n - a.s.$. In the same way $h'_\mu = \tilde{h}_\mu$, $\eta_n - a.s.$. Hence h_μ 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

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

Consequently μ is the measure $h_\mu \eta_n$. For the last assertion, we see that

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

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 converges to f over E_n, and $\mathbb{E} V_n(f_k)^2 \to \mathbb{E} V_n(f)^2$ by the expression for $\mathbb{E} V_n(f)^2$.

4.2. Some contraction properties

By the contraction principle, the moderate deviation principles presented in Theorem 2 can be transferred to continuous transformations of the local sampling random fields V_n^N.

Proposition 4. *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*

$$\begin{aligned} J_n(\nu) &= \inf \Big\{ \sum_{p=0}^n I_p(\mu_p) : (\mu_p)_{0 \le p \le n} \in \prod_{p=0}^n \mathcal{M}(E_p) \text{ s.t. } \nu = \sum_{p=0}^n \mu_p \mathcal{D}_{p,n} \Big\} \\ &= \sup_{f \in \mathcal{B}(E_n)} \left(\nu(f) - \frac{1}{2} \mathbb{E} \left(W_n(f)^2 \right) \right). \end{aligned} \tag{4.3}$$

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 (4.3) is an immediate consequence of Theorem 2. On the other hand, using (2.14) and (2.15), we prove that the random sequences

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

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

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

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

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

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

The last formula comes from the uniqueness property of the rate function, and ends the proof.

5. Moderate Deviations for Stochastic Processes

This section is mainly concerned with the proof of Theorem 3. By a recent theorem of Arcones (2003b), this result is a direct consequence of the following lemma.

Lemma 6. *Under the conditions* **(A1)** *and* **(A2)**, *for any* $y > 0$ *we have*

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

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

Proof. Consider a collection of independent random variables $X = (X^i)_{i\geq 1}$, with respective distributions $(\mu^i)_{i\geq 1}$ on some measurable state space $(E, \mathcal{E})$, with $\mathcal{F}$ a given collection of measurable functions on E s.t. $\|f\| \leq 1$.

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 $\inf_\emptyset = \infty$. Take $V(X) = \sqrt{N}\ (m(X) - \mu)$ as the fluctuation random field associated with the empirical measures $m(X) = (1/N)\sum_{i=1}^N \delta_{X^i}$ and their mean average $\mu = (1/N)\sum_{i=1}^N \mu^i$. Then

$$\pi_\psi\left(\|V(X)\|_{\mathcal{F}}\right) \leq c\ I(\mathcal{F}) \tag{5.1}$$

for some finite constant $c < \infty$. We further assume that

$$\sqrt{N}\ \pi_\psi\left(\|\mu - \overline{\mu}\|_{\mathcal{F}}\right) \leq \tau(I(\mathcal{F})) \tag{5.2}$$

some probability measure $\overline{\mu}$ on E and some non decreasing function τ. In this situation, we have

$$\mathbb{E}\left(e^{t\|V(X)\|_{\mathcal{F}(\delta)}}\right) \leq 4\ \exp\left(\frac{t^2}{2}\left[a_\delta(\mathcal{F})^2 + \frac{1}{N}\ (tb_\delta(\mathcal{F}))^2\right]\right) \tag{5.3}$$

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

$$a_\delta(\mathcal{F}) \leq c\int_0^\delta \sqrt{\log\mathcal{N}(\mathcal{F},\epsilon)}\, d\epsilon \quad \text{and} \quad b_\delta(\mathcal{F}) \leq c\ \log\mathcal{N}(\mathcal{F},\delta)\ [I(\mathcal{F}) + \tau(c\ I(\mathcal{F}))]$$

and some finite constant $c < \infty$. These results are known, so their proofs are omitted. Proofs are in Appendix C of the Web Appendix.

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

$$I(\Sigma_n(\mathcal{F}_{n+1})) < c'(n) \ I(\mathcal{F}_{n+1}) \tag{5.4}$$

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

$$\sqrt{N} \ \left\|\Phi_{n+1}(\eta_n^N) - \Phi_{n+1}(\eta_n)\right\|_{\mathcal{F}_{n+1}} \leq c(n) \ \left\|W_n^N\right\|_{\Sigma_n(\mathcal{F}_{n+1})}. \tag{5.5}$$

On the other hand, we have

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

and therefore

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

with $\Sigma_{p,n} = \Sigma_p \circ \Sigma_{p+1,n}$, and $c_p(n) = \prod_{p \leq q < n} c(q)$. From previous calculations, we have

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

Combining (5.1) with (5.4), we find that

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

for some finite constants $c''(n)$. By (5.5), we also have that

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

for some finite constants $c'''(n)$. This shows that the random fields V_n^N satisfy (5.2).

Arguing as above, we prove that

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

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 $\beta_p(n) < \infty$. It follows that

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

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 (5.3) we readily check that

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

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

$$a_\delta(\mathcal{F}) \leq c \int_0^\delta \sqrt{\log \mathcal{N}(\mathcal{F}, \epsilon)}\ d\epsilon \to_{\delta\downarrow 0}\ 0 \Rightarrow -\frac{y_{p,n}^2}{2a_{(\beta_p(n)\delta)}(\mathcal{F}_{p,n})^2} \to_{\delta\downarrow 0}\ -\infty.$$

This ends the proof of the lemma.

Acknowledgement

This work is supported by NSFC (No. 11201487).

References

Arcones, M. A. (2003a). The large deviation principle for stochastic processes I and II. *Theory of Probability and its Applications* **47**, 567-583 and **48**, 19-44.

Arcones, M. A. (2003b). Moderate deviations of empirical processes. Stochastic inequalities and applications. *Progr. Probab* **56**, 189-212, Birkhauser, Basel.

Assaraf, R. and Caffarel, M. (2000). *A Pedagogical Introduction to Quantum Monte Carlo. Mathematical models and methods for ab initio Quantum Chemistry*, 45-73 Springer doi: 10.1007/978-3-642-57237-1_3.

Assaraf, R., Caffarel, M. and Khelif, A. (2000). Diffusion Monte Carlo with a fixed number of walkers. Physical Revue E.

Assaraf, R., Caffarel, M. and Khelif, A. (2007). The Fermion Monte Carlo revisited. *J. Phys. A : Math. Theor.* **40**, 1181.

Brooks, J. K. (1969). Representations of weak and strong integrals in Banach spaces. *Proc. Nat. Acad. Sci. U.S.A.* **63**, 266-270.

Caffarel, M. (1989). Stochastic methods in quantum mechanics. In *Numerical Determination of the Electronic Structure of Atoms, Diatomic and Polyatomic Molecules*, Kluwer Academic Publishers, 85-105.

Caffarel, M. (2011). *Encyclopedia of Applied and Computational Mathematics*, ed. Bjorn Engquist, Springer.

Caffarel, M., Assaraf, R., Khelif, A., Scemama, A. and Ramirez-Solis, A. (2006). A few aspects of QMC for molecules. In *Mathematical and Numerical Aspects of Quantum Chemistry Problems, Mathematisches Forschunginstitut Oberwolfach*, p.7 Report No.47/2006.

Caffarel, M. and Ramirez-Solis, A. (2008). Fixed-Node Quantum Monte Carlo for Chemistry. In *The proceedings of the 14th International Conference: Recent Progress in Many-Body Theories, edited by Jordi Boronat, Gregory Astrakharchik, and Ferran Mazzanti*, World Scientific.

de Acosta, A. (1997). Moderate deviations for empirical measures of markov chains: lower bounds. *Ann. Probab.* **25**, 259-284.

de Acosta, A. and Chen, X. (1998). Moderate deviations for empirical measures of Markov chains: Upper bounds. *J. Theoret. Probab.* **11**, 1075-1110

Del Moral, P. (2004). *Feynman-Kac Formulae. Genealogical and Interacting Particle Systems with Applications*. Probability and its Applications, Springer Verlag, New York.

Del Moral, P. (2013). *Mean field simulation for Monte Carlo integration*. Chapman-Hall/CRC, Monographs on Statistics & Applied Probability.

Del Moral, P., Doucet, A. and Jasra, A. (2006). Sequential Monte Carlo Samplers. *J. Roy. Statist. Soc. Ser. B* **68**, 411-436.

Del Moral, P. and Ledoux, M. (2000). On the convergence and the applications of empirical processes for interacting particle systems and nonlinear filtering. *J. Theoret. Probab.* **13**, 225-257.

Del Moral, P. and Rio, E. (2011). Concentration inequalities for mean field particle models. *Ann. Appl. Probab.* **21**, 1017-1052.

Dobrushin, R. L. and Shlosman, S. B. (1994). Large and moderate deviations in the Ising model. *Adv. Soviet Math.* **20**, 91.

Doucet, A., de Freitas, J. F. and Gordon, N. J. (2001). *Sequential Monte-Carlo Methods in Practice*. Springer Verlag New York.

Gao, F. Q. (1996). Moderate deviations for martingales and mixing random processes. *Stochastic Process. Appl.* **61**, 263-275.

Gao, F. Q. (2003). Moderate deviations and large deviations for kernel density estimators. *J. Theoret. Probab.* **16**, 401-418.

Gao, F. Q. and Zhao, X. (2011). Delta method in large deviations and moderate deviations for estimators. *Ann. Statist.* **39**, 1211-1240.

Hetherington, J. H. (1984). Observations on the statistical iteration of matrices. *Physical Revue A* **30**, 2713-2719.

Kolokoltsov, V. N. (2007). Nonlinear Markov semigroups and interacting levy type processes. *J. Statist. Phys.* **126**, 585-642.

Kolokoltsov, V. N. (2010). *Nonlinear Markov Processes and Kinetic Equations*. Cambridge University Press.

Kolokoltsov, V. N. (2013). Nonlinear Levy and nonlinear Feller processes. In *Mathematics and Life Sciences. De Gruyter* (Edited by A. V. Antoniouk and R. V. Melnik), 45-70.

Ledoux, M. (1992). Sur les déviations modérées de sommes de variable aléatoires vectorielles indépendantes de même loi. *Ann. Inst. Henri Poincaré* **28**, 267-280.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces, Isometry and Processes.* Springer-Verlag Berlin.

Peligrad, M., Sang, H., Zhong, Y. and Wu, W. B. (2013). Exact moderate and large deviations for linear processes. ArXiv:1111.0537v1. In revision for *Statist. Sinica.*

Revuz D. (1976). *Markov Chains*, North-Holland.

Talagrand, M. (1994). Sharper bounds for Gaussian and empirical processes. *Ann. Probab.* **22**, 28-76.

van der Vaart A. W. and Wellner J. A. (1996). *Weak Convergence of Stochastic Processes.* Springer Series in Statistics, Springer.

Wu, L. M. (1994), Large deviations, moderate deviations and LIL for empirical processes. *Ann. Probab.* **22**, 17-27.

Wu, L. M. (1999), Forward-Backward martingale decomposition and compactness results for additive functionals of stationary ergodic Markov processes. *Annales de l'I.H.P., section B* **35**, 121-141.

Wu, W. B. and Zhao, Z. (2008). Moderate deviations for stationary processes. *Statist. Sinica* **18**, 769-782.

School of Mathematics and Statistics, University of New South Wales, Sydney, Australia.

E-mail: p.del-moral@unsw.edu.au

Department of Statistics, Zhongnan University of Economics and Law, Wuhan 430073, China.

E-mail: hu_shulan@yahoo.com

Laboratoire de Mathématiques Appliquées, CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubiere, France.

Institute of Applied Mathematics, Chinese Academy of Sciences, 100190 Beijing, China.

E-mail: li-ming.wu@math.univ-bpclermont.fr; wuliming@amt.ac.cn

(Received May 2013; accepted May 2014)