Chapter 1

LARGE DEVIATIONS FOR INTERACTING PROCESSES IN THE STRONG TOPOLOGY

Donald A. Dawson
School of Mathematics and Statistics,
Carleton University 1125 Colonel By Drive Ottawa, Canada, K1S 5B6
ddawson@math.carleton.ca

Pierre Del Moral
Lab. Stat. Probab. C.N.R.S. UMR5583
University Paul Sabatier, 118 Rt. de Narbonne, Toulouse 31062, France
delmoral@cict.fr

Abstract Strong large deviations principles for a general class of discrete generation and interacting particle systems are developed. The analysis is essentially conducted through an original projective interpretation of the τ-topology, combined with a powerful projective transfer result presented by the first author with J. Gärtner. These large deviations principles simplify and encompass the ones obtained in an earlier joint work of the second author with A. Guionnet. They are illustrated with simplified versions of McKean-Vlasov diffusions, and Boltzmann type collision models. We also describe the impact of this analysis on a recently developed class of genealogical and interacting particle interpretations of non linear Feynman-Kac-Schrödinger path measures.

1. Introduction

1.1 Description of the models

In this article, we design a natural and general projective technique for proving strong large deviations principles, for a general class of interacting particle systems. This class of interacting measure valued processes arise in various scientific disciplines, including in biology and physics, but also in diverse branches of engineering science. In order that this article is both broad, and has some theoretical and applied basis, we shall briefly review the impact of these deviations on simplified McKean-Vlasov diffusions and Boltzmann type collision models, as well as on a general and abstract class of Feynman-Kac-Schrödinger semigroups. Besides their important physical aspect, these models also arise in the analysis of nonlinear filtering, Bayesian statistics, rare events estimation, and directed polymers simulations. Readers who are interested in more detailed and precise applications should consult the series of complementary articles Del Moral (1998); Del Moral and Miclo (2000); Del Moral and Miclo (2001), as well as the book edited by Doucet and De Freitas and Gordon (2001).

To describe with some precision our models, we need to introduce some notation. We let $\mathbf{P}(E)$ be the set of additive and $[0,1]$-valued set functions on an Hausdorff topological space E, with the Borel σ-field $\mathcal{E}$, and equipped with the τ_1-topology of setwise convergence. Without further mention, we shall assume that $\mathbf{P}(E)$ is furnished with a σ-algebra which contains the Borel σ-field associated to the τ_1-topology on $\mathbf{P}(E)$. As usual, we denote by $\mathcal{B}_b(E)$ the Banach space of all bounded, and $\mathcal{E}$-measurable functions f on E, endowed with the supremum norm $\|f\| = \sup_{x\in E}|f(x)|$. By $\mathcal{P}(E)$ we denote the subset of probability measures on E, equipped with the relative τ-topology.

We consider a collection of Hausdorff topological spaces $(E_n,\mathcal{E}_n)_{n\geq 0}$, and a collection of continuous mappings Φ_n from $\mathbf{P}(E_{n-1})$ into $\mathbf{P}(E_n)$, such that $\Phi_n(\mathcal{P}(E_{n-1}))\subset\mathcal{P}(E_n)$. We also let η_n be the solution of the equation

$$\eta_n = \Phi_n(\eta_{n-1})\ , \quad \text{with} \quad \eta_0\in\mathcal{P}(E_0) \tag{1.1}$$

In practice, the above continuity conditions can be easily checked by recalling that $\mathbf{P}(E)$ can be represented as a subset of the algebraic dual of $\mathcal{B}_b(E)$ (see for instance Theorem C3 page 315 in Dembo and Zeitouni (1993)), and the τ_1-topology coincides with the $\mathcal{B}_b(E)$-topology on $\mathbf{P}(E)$.

The interacting particle interpretations of the equations (1.1) can be thought as a stochastic linearization technique. They are defined as a Markov chain $\xi_n^{(N)}$ taking values in the product spaces E_n^N with initial distribution $\eta_0^{\otimes N}$. Its elementary transitions $\xi_{n-1}^{(N)}\rightsquigarrow\xi_n^{(N)}$ are given, for any $n\geq 1$, by the formula

$$\mathrm{Prob}(\xi_n^{(N)}\in d(x_n^1,\ldots,x_n^N)\mid\xi_{n-1}^{(N)}) = \prod_{i=1}^N \Phi_n(m_N(\xi_{n-1}^{(N)}))(dx_n^i) \tag{1.2}$$

In the above display, $m_N : x\in E^N\to m_N(x)\in\mathbf{P}(E)$, is the mapping which associates to each configuration $x=(x^i)_{1\leq i\leq N}\in E^N$ the empirical measure $m_N(x) = \frac{1}{N}\sum_{i=1}\delta_{x^i}$, with δ_{x^i} denoting the Dirac measure at the site x^i. In other words, the configuration of the system $\xi_n^{(N)} = (\xi_n^{(N,i)})_{1\leq i\leq N}$ at time n consists in N-conditionally independent random variables, with common distribution $\Phi_n(m_N(\xi_{n-1}^{(N)}))$. We complete this short presentation with an alternative particle interpretation of (1.1). Suppose the one step mapping Φ_n can be rewritten in the following form

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

for some (non unique) collection of Markov kernels $K_{n,\eta_{n-1}}$ from E_{n-1} into E_n. In this situation, it is often more judicious to associate to (1.1) the particle approximating model defined as above, replacing (1.2) by the elementary transitions

$$\mathrm{Prob}(\xi_n^{(N)} \in d(x_n^1,\dots,x_n^N) \mid \xi_{n-1}^{(N)}) = \prod_{i=1}^N K_{n,m_N(\xi_{n-1}^{(N)})}(\xi_{n-1}^i, dx_n^i) \quad (1.3)$$

The asymptotic analysis of (1.3) is a little more tricky than the one of the models discussed in the former article. Although the deviations of (1.3) do not fit into our plot, we believe and conjecture that they can be derived using similar arguments. Finally, it is worth mentioning that the rate function which governs the large deviations of (1.3) is often larger than the one associated to (1.2), see for instance Del Moral and Zajic (2003).

1.2 Statement of some results

Before to describe the main results developed in this article, we first present a brief discussion on exponential deviation estimates. We also take this opportunity to fix some classical notation we shall currently use in the further development of this article. We recall that a Markov transition M_n, from a measurable space $(E_{n-1},\mathcal{E}_{n-1})$ into another $(E_n,\mathcal{E}_n)$, generates two integral operations

$$M_n : \mathcal{B}_b(E_n) \to \mathcal{B}_b(E_{n-1}) \quad \text{and} \quad M_n : \mathcal{P}(E_{n-1}) \to \mathcal{P}(E_n)$$

These two operations are defined, for any $(x_{n-1}, f_n) \in (E_{n-1}\times\mathcal{B}_b(E_n))$, and $(\mu_{n-1}, A_n) \in (\mathcal{P}(E_{n-1})\times\mathcal{E}_n)$ by the formulae

$$\begin{aligned} M_n(f_n)(x_{n-1}) &= \int_{E_n} M_n(x_{n-1}, dx_n)\ f(x_n) \\ \mu_{n-1}M_n(A_n) &= \int_{E_{n-1}} \mu_{n-1}(dx_{n-1})\ M_n(x_{n-1}, A_n) \end{aligned}$$

with $M_n(x_{n-1}, A_n) = M_n(1_{A_n})(x_{n-1})$. For any function $f_n \in \mathcal{B}_b(E_n)$, and any distribution $\mu_n \in \mathcal{P}(E_n)$, we often use the notation

$$\mu_n(f_n) = \int \mu_n(dx_n)\ f_n(x_n)$$

Under appropriate regularity conditions, and for both models (1.2) and (1.3), it is known that, in some sense, the particle occupation measures

$$\eta_n^N = m_N(\xi_n^{(N)})$$

converge as N tends to infinity to the desired distribution η_n. For instance, under appropriate Lipschitz-type conditions on Φ_n, we have the sub-Gaussian exponential estimate

$$\mathbb{P}(|\eta_n^N(f_n) - \eta_n(f_n)| > \epsilon) \leq c_n\ e^{-N\varepsilon^2/c_n}$$

for any $\varepsilon > 0$, and $f_n \in \mathcal{B}_n(E_n)$ with $\|f_n\| \leq 1$, and for some finite constant $c_n < \infty$ (see for instance Del Moral (1998)). As an aside, by the Borel-Cantelli Lemma we have the almost sure convergence

$$\lim_{N\to\infty} \eta_n^N(f_n) = \eta_n(f_n)$$

In addition, if we let $A_{n,\varepsilon}(f_n)$ be the τ-closed subset given by

$$A_{n,\varepsilon}(f_n) = \{\mu_n \in \mathcal{P}(E_n)\ :\ |\mu_n(f_n) - \eta_n(f_n)| > \epsilon\}$$

then we have the asymptotic deviation estimate

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(\eta_n^N \in A_{n,\varepsilon}(f_n)) \leq -J_n(A_{n,\varepsilon}(f_n))\ (\leq -\varepsilon^2/c_n)$$

for some constant $J_n(A_{n,\varepsilon}(f_n))$.

The large deviations analysis is concerned with estimating more precisely these exponential deviations rates. More generally, they provide sharp exponential estimates of the deviations of the approximating flow $(\eta_p^N)_{p\leq n}$ around the solution $(\eta_p)_{p\leq n}$ of the limiting measure valued equation (1.1). The study of the deviations of η_n^N around η_n have been started in Del Moral and Guionnet (1998) when the distribution space $\mathcal{P}(E_n)$ is equipped with the weak topology. Under appropriate regularity conditions on the mappings Φ_n, these large deviations principles were further extended in Del Moral and Zajic (2003) to path space models. We also mention that in Dawson and Feng (2001) a projective limit approach was also used to obtain the large deviations principles with good rate function on $\mathcal{P}([0,1])$, for the genetic Fleming-Viot process with state-independent mutation in the limit, as the effective population goes to infinity. However, in contrast to the Sanov case, in the Fleming-Viot case there is not a good rate function in the τ_1-topology but only in the weak topology.

The objective of the former article is to simplify and further extend these large deviations principles to the particle density profiles $(\eta_p^N)_{0\leq p\leq n}$ on the product space $\mathbf{P}^n(E) =_{\text{def.}} \prod_{p=0}^n \mathbf{P}(E_p)$ equipped with the τ_1-topology. We also assume that the mappings Φ_n satisfy the following regularity condition

(H) For any $n \geq 1$, there exists a reference measure $\lambda_n \in \mathcal{P}(E_n)$ and some $\rho_n > 0$, such that for any $\mu_{n-1} \in \mathcal{P}(E_{n-1})$ we have

$$\rho_n \; \lambda_n \leq \Phi_n(\mu_{n-1}) \ll \lambda_n$$

For sake of completeness, we recall that a sequence of measures Q^N on an Hausdorff topological space E, equipped with a σ-field $\sigma(E)$ which contains the Borel σ-field $B(E)$, satisfies the large deviations principles with rate function J, if the following conditions are met. The function $J : E \to [0,\infty]$ is lower semi-continuous, and for any pair of closed/open subsets (A, B) we have

$$\limsup_{N\to\infty} \frac{1}{N} \log Q^N(A) \leq -J(A) \quad \text{and} \quad \liminf_{N\to\infty} \frac{1}{N} \log Q^N(B) \geq -J(B)$$

In the above displayed formulae, we have used the notation $J(A) = \inf_A J$. A lower semincontinuous function J with compact level sets $J^{-1}([0,a])$, $a \in [0,\infty)$, is said to be a good rate function.

Our main result basically reads

THEOREM 1.1 *For any time horizon $n \geq 0$ the law Q_n^N of the N-particle density profiles $(\eta_p^N)_{0\leq p\leq n}$ satisfies the large deviations principles on the Cartesian product $\mathbf{P}^n(E)$ equipped with the τ_1-topology with good rate function*

$$J_n((\mu_p)_{0\leq p\leq n}) \;\; = \;\; \mathrm{Ent}(\mu_0|\eta_0) + \sum_{p=1}^n \; \mathrm{Ent}(\mu_p|\Phi_p(\mu_{p-1}))$$

where $\mathrm{Ent}(\,.\,|\,.\,)$ *stands for the traditional Boltzmann or the Shannon-Kullback entropy criterium, with the convention $J_n(\mu) = \infty$ when the flow $\mu = (\mu_p)_{0\leq p\leq n} \notin \mathcal{P}^n(E) =_{\text{def.}} \prod_{p=0}^n \mathcal{P}(E_p)$.*

As we shall see, the above result implies the large deviations principles for the τ-topology (and hence the weak topology) on $\mathcal{P}^n(E)$ with good rate function J_n. Also notice that $J_n(\mu) = 0$ if and only if the flow $\mu = (\mu_p)_{0\leq p\leq n}$ satisfies the non linear limiting equation (1.1) (with $\mu_0 = \eta_0$).

At first guess, the additive structure of the rate function J_n seems to indicate that the local sampling errors induced at each elementary steps of the particle approximation model propagate with respect to the time parameter. On the other hand we also know from Del Moral and Miclo (2000) that it is possible to control uniformly the propagation of the $\mathbb{L}_p$-errors with respect to the time parameter as soon as the limiting equation (1.1) is sufficiently stable.

To get one step further in our discussion, we need to introduce some additional notation. We let $\|. - .\|_{tv}$ be the total variation distance between probabilities defined by

$$\begin{aligned}\|\mu - \nu\|_{tv} &= \sup_{A\in\mathcal{E}} |\mu(A) - \nu(A)| \\ &= 2^{-1} \sup \{|\mu(f) - \nu(f)| \ : \ f \in \mathcal{B}_b(E), \quad \text{with} \quad \|f\| \leq 1\}\end{aligned}$$

We also associate to a finite partition $U = (U^i)_{1\leq i\leq d}$ of the state space E, the Kolmogorov-Smirnov seminorm on $\mathcal{P}(E)$ defined by

$$d_U(\mu, \nu) = \frac{1}{2}\sum_{i=1} |\mu(U^i) - \nu(U^i)|$$

Finally, we introduce the following regularity and Lipschitz type condition on the semi-group $\Phi_{p,n}(\eta_p) = \eta_n$ of the limiting model (1.1)

(L) For any $p \leq n$ there exists some positive constants $\beta(\Phi_{p,n})$ such that $\beta(\Phi) =_{def.} \sup_{n\geq 0} \sum_{p\leq n} \beta(\Phi_{p,n}) < \infty$ and for any pair of measures $(\mu_p, \nu_p) \in \mathcal{P}(E_p)$

$$\|\Phi_{p,n}(\mu_p) - \Phi_{p,n}(\nu_p)\|_{tv} \ \leq \ \beta(\Phi_{p,n}) \ \|\mu_p - \nu_p\|_{tv} \tag{1.4}$$

Corollary 1.2 *Assume that the regularity condition (L) is met for some finite constant $\beta(\Phi) < \infty$. Then, for any $\varepsilon > 0$, and for any finite Borel partitions $U_n = (U_n^i)_{1\leq i\leq d_n}$ of E_n, we have the uniform asymptotic estimate*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}\left(\sup_{p\leq n} d_{U_p}(\eta_p^N, \eta_p) \geq \varepsilon\right) \leq -\varepsilon^2/\beta(\Phi)$$

The article is organized as follows:

In section 1.2, we illustrate the large deviations analysis developed in this article on three different classes of interacting and measure valued processes. The first two models are more related to particle physics and fluid mechanics. They can be interpreted as simplified and discretized versions of traditional McKean-Vlasov diffusions, and Boltzmann collision models. The third and last application is concerned with an abstract class of Feynman-Kac-Schrödinger semigroups and their genealogical and interacting particle interpretation. The initiated reader can skip this section.

In section 1.3, we revisit the projective limit approach to large deviations introduced in a joint work of the first author with J. Gärtner, c.f. Dawson and Gärtner (1987), and further developed in a series of three articles de Acosta (1994a); de Acosta (1994b); de Acosta (1997).

We provide an original min-max type theorem, and an explicit formula expressing good rate functions in terms of their image measures, with respect to the projection coordinate mappings.

In section 1.4, we show that $\mathbf{P}(E)$, furnished with the τ_1-topology, is homeomorphic to a compact projective limit of finite dimensional distribution spaces. We combine this interpretation with the projective contraction results derived in section 1.3, to show that the finite space or the strong version of Sanov's Theorem on an Hausdorff topological space are in fact equivalent.

Section 1.5 is mainly concerned with the proof of Theorem 1.1, and Corollary 1.2. We first extend the projective limit interpretation of the τ_1-topology derived in section 1.4 to path-distribution spaces, and by a simple induction we prove that J_n are good rate functions on $\mathbf{P}^n(E)$.

Our approach to Theorem 1.1 is roughly as follows. The proof of the large deviations upper bound is essentially based on the min-max Theorem derived in section 1.3, while the proof of the large deviations lower bound is conducted by induction on the time parameter, in the same spirit as in the proof of the strong version of Sanov's Theorem.

2. Motivating Examples

2.1 McKean-Vlasov and Boltzmann type models

The McKean-Vlasov and the Boltzmann equations are nonlinear partial and integro-differential equations. Both equations arise in fluid mechanics, and more precisely in the microscopic and macroscopic modeling of gases. For a review on these equations, we refer the reader to the review article of Méléard (1995), and references therein. The central idea behind the particle interpretation of these models is to construct a nonhomogenous Markov process whose distribution laws are (weak) solution of the desired equation in distribution space. The existence and uniqueness of this stochastic model often requires fixed point arguments on path-space distributions, see for instance Theorem 1.1, page 172, in Sznitman (1991). In contrast to the continuous time situation, the discrete time analysis is based on elementary Markov chain techniques.

To give a brief presentation of these models, we recall that the McKean-Vlasov equations correspond to mean-field drift diffusions describing the repulsive or attractive interaction between close gas molecules. The simplified discrete time, and one dimensional version, of this model is given by the following stochastic recursion

$$X_n = a_n(X_{n-1}, \eta_{n-1}) + b_n(X_{n-1}, \eta_{n-1})\ W_n$$

where (a_n, b_n) are bounded functions from $(\mathbb{R} \times \mathcal{P}(\mathbb{R}))$ into $\mathbb{R}$, with $b_n(x, \mu) > 0$ for any $(x, \mu) \in (\mathbb{R} \times \mathcal{P}(\mathbb{R}))$, and $(W_n)_{n \geq 0}$ are independent and identically distributed Gaussian random variables with $\mathbb{E}(W_n) = 0$ and $\mathbb{E}(W_n^2) = 1$. The initial condition is defined by some random variable X_0, with a given distribution $\eta_0 \in \mathcal{P}(\mathbb{R})$.

If we let η_n be the distribution of the random state X_n at time n, then we find that the flow η_n satisfies the following nonlinear measure valued equation

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

with the Markov transition $K_{n,\eta_{n-1}}$ defined by

$$K_{n,\eta_{n-1}}(x_{n-1}, dx_n) = \frac{1}{\sqrt{2\pi}}\ \exp\left\{ -\frac{1}{2} \left(\frac{x_n - a_n(x_{n-1}, \eta_{n-1})}{b_n(x_{n-1}, \eta_{n-1})} \right)^2 \right\}\ dx_n$$

Suppose the functions $\mu \to a_n(x, \mu)$ and $\mu \to b_n(x, \mu)$ are uniformly τ-continuous with respect to the parameter x. Also assume that we have $\inf b_n^2 \geq \delta$, for some $\delta > 0$. Under these regularity conditions, we easily prove that the mappings Φ_n are τ-continuous, and they satisfy condition (H) with the constants

$$\rho_n = e^{-\frac{\|a_n\|^2}{\delta}}\ p_n(\mathbb{R})$$

and the reference measures

$$\lambda_n(dx) = p_n(dx)/p_n(\mathbb{R}) \quad \text{with} \quad p_n(dx) = \exp\left(-\delta^{-1}\ [y^2/2 + \|a_n\|\ |y|] \right)$$

The N-particle model (1.3) associated to the McKean transitions $K_{n,\eta_{n-1}}$ consists in N-interacting particles $\xi_n^{(N)} = (\xi_n^{(N,i)})_{1 \leq i \leq N}$ evolving according to the simple stochastic recursive equations, defined for any $1 \leq i \leq N$ by

$$\xi_n^{(N,i)} = a_n \left(\xi_{n-1}^{(N,i)}, \frac{1}{N} \sum_{j=1}^{N} \delta_{\xi_{n-1}^{(N,j)}} \right) + b_n \left(\xi_{n-1}^{(N,i)}, \frac{1}{N} \sum_{j=1}^{N} \delta_{\xi_{n-1}^{(N,j)}} \right)\ W_n^i$$

where $(W_n^i)_{1\leq i\leq N}$ are N independent copies of the sequence of Gaussian random variables W_n, and $(\xi_0^{(N,i)})_{1\leq i\leq N}$ are N independent copies of the initial random state X_0. The N-particle model (1.2) is defined in the same way, by introducing at each step an extra level of randomness.

Boltzmann equations represent the macroscopic evolution of rarefied gas models. In contrast to the McKean-Vlasov case, the interaction between gas molecules is now interpreted as a jump/collision mechanism. The simplified discrete time Boltzmann models described hereafter are rather close to the class of Feynman-Kac-Schrödinger semigroups developed in the next section.

To describe with some generality these equations, it is convenient to introduce some precise topological structure. We let $(E_n, B(E_n))_{n\geq 0}$ be a sequence of Polish spaces. We recall that a Polish space is a complete separable metric space. We denote by $\mathcal{C}_b(E)$, the set of all continuous and bounded functions on some Polish space E.

We let $g_n \in \mathcal{C}_b(E_n^2)$, be a sequence of pair-collision energy positive functions. By $Q_n((x_{n-1}, x'_{n-1}), dx_n)$ we denote a sequence of Markov transitions from E_{n-1}^2 into E_n. The quantity $g_n(x_n, x'_n)$ can be interpreted as the likelihood of interaction between a pair of particles at sites (x_n, x'_n). Given a pair of colliding sites (x_n, x'_n), the distribution $Q_n((x_{n-1}, x'_{n-1}), dx_n)$ represents both, the statistics of the resulting collision, as well as the free evolution of a particle between collisions.

We associate to the pair (g_n, Q_n) the one step mappings Φ_n defined for any $\eta_n \in \mathcal{P}(E_n)$ by

$$\Phi_{n+1}(\eta_n)(dx_{n+1}) = \int_{E_n^2} \Psi_n(\eta_n\otimes\eta_n)(d(x_n, x'_n))\ Q_{n+1}((x_n, x'_n), dx_{n+1})$$

with the Boltzmann-Gibbs measures $\Psi_n(\eta_n\otimes\eta_n) \in \mathcal{P}(E_n^2)$ associated to the collision-energy function g_n, and defined by

$$\Psi_n(\eta_n\otimes\eta_n)(d(x_n, x'_n)) = \frac{g_n(x_n, x'_n)}{(\eta_n\otimes\eta_n)(g_n)}\ (\eta_n\otimes\eta_n)(d(x_n, x'_n))$$

As usual, the corresponding N-particle model consists in N-interacting particles. The elementary transition (1.2) is decomposed into two separate mechanisms

$$\xi_n^{(N)} \in E_n^N \xrightarrow{\text{pair}-\text{collisions}} \widehat{\xi}_n^{(N)} \in (E_n^2)^N \xrightarrow{\text{free evolution}} \xi_{n+1}^{(N)} \in E_{n+1}^N$$

During the pair-collisions transition, we select randomly N pairs of colliding particles $\widehat{\xi}_n^{(N,i)} = (\widehat{\xi}_n^{(N,i,1)}, \widehat{\xi}_n^{(N,i,2)})$ with the discrete distribution

$$\Psi_n(m_N(\xi_n^{(N)}) \otimes m_N(\xi_n^{(N)}))$$

$$= \sum_{1\leq i,j\leq N} \frac{g_n(\xi_n^{(N,i)},\xi_n^{(N,j)})}{\sum_{1\leq k,l\leq N} g_n(\xi_n^{(N,k)},\xi_n^{(N,l)})}\ \delta_{(\xi_n^{(N,i)},\xi_n^{(N,j)})}$$

Notice that a pair of particles $(\xi_n^{(N,i)}, \xi_n^{(N,j)})$ is more likely to collide when its interaction-energy $g_n(\xi_n^{(N,i)}, \xi_n^{(N,j)})$ is high. In practice, close particles in some local neighborhood are more likely to collide. Given a selected pair, say

$$\widehat{\xi}_n^{(N,k)} = (\xi_n^{(N,i)}, \xi_n^{(N,j)})$$

the particle $\xi_n^{(N,i)}$ collides with $\xi_n^{(N,j)}$, and then evolves according to some free motion. These two mechanisms are performed by sampling a random variable $\xi_{n+1}^{(N,k)}$ with a distribution $Q_{n+1}((\xi_n^{(N,i)}, \xi_n^{(N,j)}), dx_{n+1})$ that depends on the selected colliding pair.

We further assume that $Q_{n+1}(\mathcal{B}_b(E_{n+1})) \subset \mathcal{C}_b(E_n^2)$. In this situation, we can check that the mappings Φ_{n+1} are continuous with respect to the τ-topologies from $\mathcal{P}(E_n)$ into $\mathcal{P}(E_{n+1})$. We finally observe that condition (H) is in fact equivalent to the following one. There exist some positive constants $\rho_n > 0$, and some reference measures $\lambda_n \in \mathcal{P}(E_n)$, such that for any pairs $(x_{n-1}, x'_{n-1}) \in E_{n-1}$, we have

$$\rho_n\ \lambda_n \leq Q_n((x_{n-1}, x'_{n-1}), .) \ll \lambda_n$$

To give a physical sound to these abstract models, we examine the situation where $\eta_n(dx_n)$ represents the distribution of particle

$$x_n = (p_n, v_n) \in E_n = (\mathbb{R}^3 \times \mathbb{R}^3)$$

in position $p_n \in \mathbb{R}^3$, with speed $v_n \in \mathbb{R}^3$. The collision energy function $g_n(x_n, x'_n)$ between two particles $x_n = (p_n, v_n)$ and $x'_n = (p'_n, v'_n)$ is often expressed by a cross-section function that depends on the pair positions (p_n, p'_n), and only on the quantities $|v_n, v'_n|$, and $(v_n - v'_n).u$, where u is a given unit vector. For a more precise description of these cross-sections, we refer the reader to the review article of Méléard (1995), and references therein. The prototype of kernel Q_{n+1} is defined as the composition of a binary collision-speed exchange transition S_n, with a free evolution kernel M_{n+1}

$$Q_{n+1}((x_n, x'_n), dx_{n+1}) = \int\ S_n((x_n, x'_n), d(y_n, y'_n))\ M_{n+1}(y'_n, dx_{n+1})$$

The form of S_n is dictated by the conservation of kinetic energy, and momentum for binary collisions. It is defined for any $x_n = (p_n, v_n)$ and $x'_n = (p'_n, v'_n)$ by the formula

$$S_n((x_n, x'_n), .) = \delta_{([p_n, v_n + ((v'_n - v_n).u)\ u], [p'_n, v'_n + ((v_n - v'_n).u)\ u])}$$

Notice that this simplified Boltzmann transition Q_{n+1} describes precisely the collision, and the free motion of the first component particle, but it doesn't says anything on the colliding partner. This question can be answered by a traditional state space enlargement. Finally and in this situation, we observe that the regularity condition (H) is met, as soon as the free motion Markov transitions M_n satisfy this condition. That is, when we have

$$\rho_n\ \lambda_n \leq M_n(x_{n-1}, .) \ll \lambda_n$$

For instance, arguing as in the McKean-Vlasov case, this condition is met for regular Gaussian transitions.

2.2 Feynman-Kac-Schrödinger semigroups

As we mentioned in the introduction, these nonlinear semigroups are at the corner of a variety of application areas. Their probabilistic interpretations depends on the physical, biological, or engineering problem at hand. To capture, both their generality, and their full impact in applications, it is convenient to introduce these models using a rather abstract definition.

We let $(E_n, \mathcal{E}_n)_{n\geq 0}$ be a sequence of Hausdorff topological spaces. We consider a sequence of positive potential functions $G_n \in \mathcal{B}_b(E_n)$, and a collection of Markov transitions $M_n(x_{n-1}, dx_n)$ from E_{n-1} into E_n. We associate to the pair potential/transition (G_n, M_n), the distribution flow η_n defined by the Feynman-Kac measures

$$\eta_n(f_n) = \gamma_n(f_n)/\gamma_n(1) \quad \text{with} \quad \gamma_n(f_n) = \mathbb{E}\left(f_n(X_n) \prod_{0\leq p<n} G_p(X_p)\right)$$

where X_n is an E_n-valued Markov chain with some initial distribution $\eta_0 \in \mathcal{P}(E_0)$, and elementary transitions M_n. Using the multiplicative structure, and the Markov property, we readily check that the flow η_n satisfies the following equation

$$\eta_n = \Phi_n(\eta_{n-1}) =_{\text{def.}} \Psi_{n-1}(\eta_{n-1})M_n$$

with the Boltzmann-Gibbs transformation Ψ_n on $\mathcal{P}(E_n)$ defined by

$$\Psi_n(\eta_n) = \frac{1}{\eta_n(G_n)}\ G_n(x_n)\ \eta_n(dx_n)$$

Since for any $f \in \mathcal{B}_b(E_{n+1})$ we have

$$\Phi_{n+1}(\eta_n)(f) = \eta_n(G_n M_{n+1}(f))/\eta_n(G_n)$$

we readily check that Φ_{n+1} is a τ_1-continuous mapping from $\mathbf{P}(E_{n-1})$ into $\mathbf{P}(E_n)$ and $\Phi_n(\mathcal{P}(E_{n-1})) \subset \mathcal{P}(E_n)$. Also notice that condition (H) is equivalent as to have, for any $x_{n-1} \in E_{n-1}$, some positive constants $\rho_n > 0$, and some reference measures $\lambda_n \in \mathcal{P}(E_n)$

$$\rho_n \lambda_n \leq M_n(x_{n-1}, .) \ll \lambda_n$$

For instance, let us suppose that $E_n = \mathbb{R}$ and

$$M_n(x, dy) = \frac{1}{\sqrt{2\pi}}\; e^{-\frac{1}{2}(y-a_n(x))^2}\; dy$$

where a_n is a bounded measurable drift function on $\mathbb{R}$. In this case, condition (H) is clearly met, with the parameters $\rho_n = e^{-\|a_n\|^2/2}\; p_n(\mathbb{R})$, and the reference measures defined by

$$\lambda_n(dx) = p_n(dx)/p_n(\mathbb{R}) \quad \text{with} \quad p_n(dx) = \exp\left(-[y^2/2 + \|a_n\|\; |y|]\right)$$

In connection with Corollary 1.2, we mention the Feynman-Kac semigroup $\Phi_{p,n}$ satisfies condition (L) as soon as the pair potential/kernel (G_n, M_n) are such that for any $(x_n, y_n) \in E_n^2$ and $n \geq 0$

$$G_n(x_n) \geq \varepsilon(G)\; G_n(y_n)$$

and

$$M_{n,n+m}(x_n, .) \geq \varepsilon(M)\; M_{n,n+m}(y_n, .)$$

for some $m \geq 1$ and some pair of constants $(\varepsilon(G), \varepsilon(M)) \in (0,1]^2$. In the above display, $M_{n,n+m} = (M_{n+1} \ldots M_{n+m})$ stands for the composition of the Markov kernels M_p from $p = (n+1)$ to $p = (n+m)$. In this situation, it is known that (1.4) is met with

$$\beta(\Phi_{p,n}) = \frac{2}{\varepsilon(M)\varepsilon(G)^m}\; (1 - \varepsilon(M)^2 \varepsilon(G)^{m-1})^{[(n-p)/m]}$$

We conclude that the estimate stated in Corollary 1.2 holds true with

$$\beta(\Phi) \leq 2m/(\varepsilon(M)^3 \varepsilon(G)^{(2m-1)})$$

For more details we refer the reader to Del Moral and Miclo (2000) and references therein.

Notice that the particle interpretation (1.2) of the Feynman-Kac-Schrödinger flow is a simple selection/mutation genetic algorithm. To be more precise, as in the simplified Boltzmann type model described in the previous section, its elementary transitions (1.2) are decomposed into two separate mechanisms

$$\xi_n^{(N)} \in E_n^N \xrightarrow{\text{selection}} \widehat{\xi}_n^{(N)} \in E_n^N \xrightarrow{\text{mutation}} \xi_{n+1}^{(N)} \in E_{n+1}^N$$

During the first transition, we select randomly N particles $\widehat{\xi}_n^{(N,i)}$ with the discrete distribution

$$\Psi_n(m_N(\xi_n^{(N)})) = \sum_{1\leq j\leq N} \frac{G_n(\xi_n^{(N,j)})}{\sum_{1\leq k\leq N} G_n(\xi_n^{(N,k)})}\, \delta_{\xi_n^{(N,j)}}$$

Then, we evolve each selected particle $\widehat{\xi}_n^{(N,i)}$, with an elementary move $\widehat{\xi}_n^{(N,i)} \rightsquigarrow \xi_{n+1}^{(N,i)}$, according to the transition distribution $M_{n+1}(\widehat{\xi}_n^{(N,i)}, .)$.

A final remark is that the random variables X_n may represent the path of an auxiliary Markov chain X'_p, from the origin up to time n, and taking values in some Hausdorff topological spaces E'_p. That is, we have that

$$X_n = (X'_0, \ldots, X'_n) \in E_n = (E'_0 \times \ldots \times E'_n) \tag{1.5}$$

In this case, the corresponding particle interpretation model forms a genetic algorithm taking values in path space. The mutation mechanism consists in extending each path particle with an elementary transitions. Notice that the selection pressure now depends on the path-potential. Whenever the potential functions only depend on terminal values, this path-particle model can be interpreted as a genealogical tree evolution model (see Del Moral and Miclo (2001) and references therein). In this situation, the Feynman-Kac-Schrödinger model doesn't satisfy anymore the regularity condition (H), and the large deviations analysis developed in this article only provides a large deviation upper bound.

We illustrate this abstract class of semigroups with an elementary filtering problem. We let X_n be an $\mathbb{R}$-valued Markov chain, with initial distribution $\eta_0 = \text{Law}(X_0)$, and elementary Markov transitions M_n. We also let $(Y_n)_{n\geq 0}$, be the random observation sequence defined by

$$Y_n = h_n(X_n) + V_n$$

The sensor function h_n is assumed to be measurable and bounded, and the noise disturbances V_n are independent, and identically distributed, Gaussian random variables with $\mathbb{E}(V_n) = 0$, and $\mathbb{E}(V_n^2) = 1$. To get one step further, we fix a sequence of observations $Y_n = y_n$, $n \geq 0$, and we consider the nonhomogeneous potential functions G_n defined by

$$G_n(x) = \exp\left\{-\frac{1}{2}\,(y_n - h_n(x))^2\right\}$$

In this situation, the Feynman-Kac flow η_n associated to the pair potential/transition (G_n, M_n) is a version of the conditional distribution of X_n given $Y_p = y_p$, $p < n$. That is, we have that

$$\eta_n = \mathrm{Law}(X_n |\ Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1})$$

If we consider the path-space model $X_n = (X'_0, \ldots, X'_n)$, then we clearly have that

$$\eta_n = \mathrm{Law}((X'_0, \ldots, X'_n) \mid Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1})$$

In this filtering context, the genetic particle model and its genealogical tree, can be interpreted both, as a natural learning prediction/updating process, but also as a stochastic and adaptive refining grip approximation model. During the mutation/prediction, each individual particle explores the space of possible values of the unobserved transitions of the signal. When the current observation is available, the selection process updates these predictions, and gives more reproductive opportunities to more likely predictions.

This class of Feynman-Kac-Schrödinger models in path-spaces have been extensively studied in the literature (see Del Moral and Miclo (2000); Del Moral and Miclo (2001), and references therein). They arise in various scientific disciplines. In signal processing, and Bayesian statistic literature, they represent the conditional path-distribution of a Markov signal given a sequence of partial and noisy observations. In biology, they are used to model random polymerizations of macro-molecules, in chemical solvents. In physics, they also describe the evolution of a Markov particle in an medium, with absorbing obstacles related to potential functions. More recently, they have been proved to be useful in the modeling of the biological and human learning process, see the recent article of Körding and Wolpert (2004).

These evolutionary particle algorithms have taken various botanical names in the applied literature, such as the “go with the winner”, the multi-splitting or the prune enrichment evolution model in biology, as well as the bootstrap or condensation filter, the sampling-importance-resampling, the spawning filter, and the population Monte-Carlo algorithm in signal processing. For more details on the theoretical and the applied side of this subject we refer the reader to Del Moral and Miclo (2000), and Doucet and De Freitas and Gordon (2001), and references therein.

3. Large deviations for projective limits

A projective (or inverse) spectrum $(M_U, p_{U,V})_{U \geq V}$ of a directed set $(\mathcal{U}, \leq)$ is a family $\mathcal{M} = (M_U, p_{U,V})_{U \geq V}$ of Hausdorff topological spaces M_U and continuous mappings

$$p_{U,V} : M_U \to M_V, \qquad U \geq V$$

with $p_{U,U} = Id$, and satisfying the compatibility conditions

$$U \geq V \geq W \Longrightarrow p_{U,W} = p_{U,V} \circ p_{V,W}$$

Let p_U be the canonical projection from $\prod_{U \in \mathcal{U}} M_U$ into M_U. The set

$$\lim_{\mathcal{U}} \mathcal{M} = \{m \in \prod_{U \in \mathcal{U}} M_U \ : \ \forall U \geq V \ \ p_V(m) = p_{U,V}(p_U(m))\}$$

equipped with the product topology, is called the projective (or inverse) limit space of the spectrum. We shall always assume that the sets M_U are equipped with the Borel σ-field $B(M_U)$ and the σ-field $\sigma(\lim_{\mathcal{U}} \mathcal{M})$ on the projective limit space contains all the sets $p_U^{-1}(B(M_U))$.

Lemma 1.3 *A function I is a good rate function on $\lim_{\mathcal{U}} \mathcal{M}$, if and only if, there exists a sequence of good rate functions I_U on M_U, with $U \in \mathcal{U}$, and such that*

$$I(m) = \sup_{U \in \mathcal{U}} I_U(p_U(m)) \tag{1.6}$$

Proof:
Suppose I is a good rate function on $\lim_{\mathcal{U}} \mathcal{M}$, in the sense that I is a lower semi-continuous function with compact level sets $I^{-1}([0,a])$, $a > 0$. By the contraction principle (see for instance Theorem 4.2.1 in Dembo and Zeitouni (1993)), the image set-functions $I_U = I \circ p_U^{-1}$ defined by

$$I_U(A) = \inf \{I(m) \ : \ m \in p_U^{-1}(A)\}$$

are good rate functions on M_U. In addition, by the lower semi-continuity property of I we have

$$I(m) = \sup \{I(m) \ : \ m \in p_U^{-1}(A_U), \quad A_U \text{ open} \subset M_U, \quad U \in \mathcal{U}\}$$

from which we prove that (1.6) holds true. Inversely, if I_U is a sequence of good rate function on the sets M_U with $U \in \mathcal{U}$, then the function I defined in (1.6) is a good rate function on I. The lower semi-continuity property comes from the fact that the pointwise supremum of a of lower semi-continuous functions is again lower semi-continuous.

To prove that I has compact level sets, we first observe that the compact levels sets $I_U^{-1}([0,a])$ of the functions I_U satisfy the compatibility condition $I_V^{-1}([0,a]) = p_{U,V}(I_U^{-1}([0,a]))$, for all $U \geq V$. Since projective limits are inherited by closed sets, the levels sets $I^{-1}([0,a]) = \lim_{\mathcal{U}} \mathcal{M} \cap \prod_{U \in U} I_U^{-1}([0,a])$ are the projective limits of compact sets; thus, they are compact by Tychonov's Theorem. ∎

Let us quote a min-max type theorem for good rate functions on projective limit spaces.

THEOREM 1.4 *Let I be a good rate function on the projective limit space $\lim_{\mathcal{U}} \mathcal{M}$, and let $I_U = I \circ p_U^{-1}$ be the p_U-image function of I on M_U. For any closed set $F \subset \lim_{\mathcal{U}} \mathcal{M}$, we have*

$$I(F) = \sup_{u \in \mathcal{U}} I_U(\overline{p_U(F)}) \tag{1.7}$$

where $\overline{A}$ stands for the closure of a measurable set A.

Proof:
By (1.6), we first notice that (1.7) can be rewritten as

$$I(F) = \inf_{m \in F} \sup_{U \in \mathcal{U}} I_U(p_U(m)) = \sup_{U \in \mathcal{U}} \inf_{m \in \overline{p_U(F)}} I_U(m) =_{def.} \tilde{I}(F)$$

Since we have $I_U(p_U(m)) \geq I_U(\overline{p_U(F)})$, for any $m \in F$, we deduce that $I(F) \geq \tilde{I}(F)$. To prove the reverse inequality, we assume that $\tilde{I}(F) < \infty$; otherwise, the desired bound trivially holds true. In this case, for each $\varepsilon > 0$ the sets

$$p_U^{-1}(\overline{p_U(F)}) \cap I^{-1}([0, \tilde{I}(F) + \varepsilon])$$

are non empty compact sets. The compactness property results from the fact that $p_U^{-1}(\overline{p_U(F)})$ is a closed set into the compact level sets $I^{-1}([0, \tilde{I}(F) + \varepsilon])$. On the other hand, if these sets were empty, we would be able to find some $m \in p_U^{-1}(\overline{p_U(F)})$ such that

$$I(m) \geq I(p_U^{-1}(\overline{p_U(F)})) = I_U(\overline{p_U(F)}) \geq \tilde{I}(F) + \varepsilon = \sup_{V \in \mathcal{U}} I_V(\overline{p_V(F)}) + \varepsilon$$

This clearly yields a contradiction. Finally, we observe that

$$U \geq V \Rightarrow p_U^{-1}(\overline{p_U(F)}) \subset p_V^{-1}(\overline{p_V(F)})$$

This readily yields that, for each $\varepsilon > 0$ we have

$$[\cap_{U \in \mathcal{U}} p_U^{-1}(\overline{p_U(F)})] \cap I^{-1}([0, \tilde{I}(F) + \varepsilon]) \neq \emptyset$$

Since F coincide with the projective limit of the sets $(\overline{p_U(F)})_{U\in\mathcal{U}}$, we have $F = \cap_{U\in\mathcal{U}} p_U^{-1}(\overline{p_U(F)})$. From previous considerations, for any $\varepsilon > 0$, there exists some point $m_\varepsilon \in F$, such that

$$I(F) \leq I(m_\varepsilon) \leq \tilde{I}(F) + \varepsilon$$

We end the proof of the theorem, by letting $\varepsilon \to 0$. ■

One easily derived consequence of Lemma 1.3 and Theorem 1.4, is the following corollary.

COROLLARY 1.5 *A sequence of probability measures P^N on $\lim_{\mathcal{U}} \mathcal{M}$ satisfy the large deviations principles with some good rate function I, if and only if, the sequence of all image measures $P^N \circ p_U^{-1}$ satisfy the large deviations principles on M_U, for some good rate functions I_U with $U \in \mathcal{U}$. In addition, the rate functions are given by $I_U = I \circ p_U^{-1}$, and they satisfy (1.6).*

Proof:
Assume that Q^N satisfy the large deviations principles on $\lim_{\mathcal{U}} \mathcal{M}$ for some good rate function I. By a traditional contraction argument, the image measures $Q^N \circ p_U^{-1}$ also satisfy the large deviations principles on M_U, with good rate function $I_U = I \circ p_U^{-1}$.

In the reverse situation, suppose that the sequence of measures $Q^N \circ p_U^{-1}$ satisfy the large deviations principles upper bound, with the good rate functions I_U. Since a closed set $F \subset \lim_{\mathcal{U}} \mathcal{M}$ coincides with the projective limit of the sets $(\overline{p_U(F)})_{U\in\mathcal{U}}$, we have $F = \cap_{U\in\mathcal{U}} p_U^{-1}(\overline{p_U(F)})$.

Applying the deviation upper bound for $Q^N \circ p_U^{-1}$, we conclude that

$$\begin{aligned}\limsup_{N\to\infty} \frac{1}{N} \log Q^N(F) &\leq \limsup_{N\to\infty} \frac{1}{N} \log Q^N \circ p_U^{-1}(\overline{p_U(F)}) \\ &\leq -I_U(\overline{p_U(F)})\end{aligned}$$

We complete the proof, by taking the infimum over all $U \in \mathcal{U}$, and using (1.7).

Suppose that $Q^N \circ p_U^{-1}$ satisfy the deviation lower bound, with the good rate function I_U. For any open subset $A \subset \lim_{\mathcal{U}} \mathcal{M}$, and $m \in A$, there exists an open set $B_U \in M_U$, for some $U \in \mathcal{U}$, such that

$$m \in p_U^{-1}(B_U) \ \subset \ A$$

Applying the deviation lower bound for $Q^N \circ p_U^{-1}$, we find that

$$\begin{aligned}\liminf_{N\to\infty} \frac{1}{N} \log Q^N(A) &\geq \liminf_{N\to\infty} \frac{1}{N} \log Q^N \circ p_U^{-1}(B_U) \\ &\geq -I_U(B_U) \geq -I_U(m)\end{aligned}$$

The end of the proof of the corollary is now clear. ∎

4. A projective approach to Sanov's Theorem

Let $\mathcal{U}$ be the set of all finite and Borel partitions of E. By $\sigma(U)$, we denote the σ-algebra generated by $U \in \mathcal{U}$. We let H be the relative entropy criteria on $\mathbf{P}(E)$, defined by

$$H(\mu|\nu) =_{def.} \sup_{U\in\mathcal{U}} \mathrm{Ent}_U(\mu|\nu)$$

where for each $U = (U^k)_{1\leq k\leq d} \in \mathcal{U}$, $\mathrm{Ent}_U(.|.)$ denotes the the U-relative entropy defined (with standard conventions) by the formulae

$$\mathrm{Ent}_U(\mu|\nu) = \sum_{k=1}^{d} \mu(U^k) \ \log\left(\mu(U^k)/\nu(U^k)\right)$$

with the convention $0\log 0 = 0 = 0\log(0/0)$. We let $V \leq U$ be the partial order on $\mathcal{U}$ induced by the inclusions relation $\sigma(V) \subset \sigma(U)$ between the corresponding σ-algebras. We slight abuse the notation, and we let $\mathcal{P}(U)$ and $\mathcal{B}_b(U)$ be respectively, the set of probability measures on $(E, \sigma(U))$, and the set of bounded and $\sigma(U)$-measurable functions on E. For each pair of partitions $V \leq U$, we let q_V and $q_{U,V}$ be the canonical projections from $\mathbf{P}(E)$ into $\mathcal{P}(V)$, and from $\mathcal{P}(U)$ into $\mathcal{P}(V)$. Notice that $\mathrm{Ent}_U(.|\nu)$ is finite and continuous on the compact sets $\{\eta \in \mathcal{P}(U) \ : \ \eta \ll q_U(\nu)\}$, and $\mathrm{Ent}_U(\mu|.)$ is continuous for every fixed μ.

Lemma 1.6 *The projective limit topological space* $\lim_{\mathcal{U}} \mathcal{P}$ *associated to the set* $(\mathcal{P}(U), q_{U,V})_{V\leq U}$ *is compact and homeomorphic to* $\mathbf{P}(E)$. *In addition, for any* $(\mu,\nu) \in (\mathbf{P}(E) \times \mathcal{P}(E))$, *with* $H(\mu|\nu) < \infty$, *we have* $\mu \in \mathcal{P}(E)$ *and*

$$H(\mu|\nu) = \mathrm{Ent}(\mu|\nu) \tag{1.8}$$

Proof:

The set $\mathcal{P} = ((\mathcal{P}(U), d_U), q_{U,V})_{U\geq V}$ forms a projective inverse spectrum of $\mathcal{U}$, with compact metric spaces $(\mathcal{P}(U), d_U)$, and connecting maps

$q_{U,V}$. The Hausdorff property ensures that $\lim_{\mathcal{U}} \mathcal{P}$ is a closed subset of $\prod_{U \in \mathcal{U}} \mathcal{P}(U)$, which is compact by Tychonov's Theorem. Let

$$h : \lim_{\mathcal{U}} \mathcal{P} \to \mathbf{P}(E)$$

be the mapping which associates to a point $\mu = (\mu^U)_{U \in \mathcal{U}} \in \lim_{\mathcal{U}} \mathcal{P}$ the set function $h(\mu) \in \mathbf{P}(E)$ defined for any $A \in \mathcal{E}$ by $h(\mu)(A) = \mu^U(A)$, where $U \in \mathcal{U}$ is some finite partition of E such that $A \in \sigma(U)$. Then, h is an homeomorphism between the compact spaces $\lim_{\mathcal{U}} \mathcal{P}$ and $\mathbf{P}(E)$. To prove this claim, we first check that h is well defined. This assertion is easily proved by noting that

$$A \in (\sigma(U) \cap \sigma(V)) \Longrightarrow A \in \sigma(U \vee V)$$

where

$$(U \vee V) = \{A \cap B \ : \ (A, B) \in (U \times V)\}$$

is the smallest partition such that $\sigma(U) \vee \sigma(V) = \sigma(U \vee V)$. By the compatibility conditions in the definition of $\lim_{\mathcal{U}} \mathcal{P}$, we have

$$\mu^{(U \vee V)}(A) = \mu^U(A) = \mu^V(A)$$

To prove that $h(\mu)$ is an additive set function on $\mathcal{E}$, we choose a pair of disjoint Borel sets A, B, and a pair U, V of partitions with $A \in U$, and $B \in V$. Since A and B are disjoint, we have

$$\begin{aligned} A &= A \cap (E - B) = A \cap (\cup_{C \in V, C \neq B} C) \\ &= \cup_{C \in V, C \neq B} (A \cap C) \in \sigma(U \vee V) \end{aligned}$$

and by symmetry arguments, $B \in \sigma(U \vee V)$. Using the fact that $\mu^{(U \vee V)} \in \mathcal{P}(U \vee V)$, we conclude that

$$\begin{aligned} h(\mu)(A \cup B) &= \mu^{(U \vee V)}(A \cup B) = \mu^{(U \vee V)}(A) + \mu^{(U \vee V)}(B) \\ &= h(\mu)(A) + h(\mu)(B) \end{aligned}$$

Let us prove that h is an injection. Let $\mu, \nu \in \lim_{\mathcal{U}} \mathcal{P}$ be a pair of points such that $h(\mu) = h(\nu)$. By definition of h, we find that $\mu^U = \nu^U$ for any $U \in \mathcal{U}$, from which we conclude that $\mu = \nu$. On the other hand, for any $\mu \in \mathbf{P}(E)$ we have $(q_U(\mu))_{U \in \mathcal{U}} \in \lim_{\mathcal{U}} \mathcal{P}$, and

$$h((q_U(\mu))_{U \in \mathcal{U}})(A) = q_U(\mu)(A) = \mu \circ e_U^{-1}(A) = \mu(A)$$

for any $A \in \sigma(U)$, for some $U \in \mathcal{U}$. We conclude that h is a bijective map from $\lim_{\mathcal{U}} \mathcal{P}$ into $\mathbf{P}(E)$, and $h^{-1}(\mu) = (q_U(\mu))_{U \in \mathcal{U}}$.

It remains to prove that h and h^{-1} are continuous. To prove this final step, we observe that for any sequence $\mu = (\mu^U)_{U \in \mathcal{U}}$ of points in $\lim_{\mathcal{U}} \mathcal{P}$, and $\mu = (\mu^U)_{U\in\mathcal{U}} \in \lim_{\mathcal{U}} \mathcal{P}$, we have the following series of equivalent assertions

$$\begin{aligned}\lim_{n\to\infty} \mu = \mu \text{ in } \lim_{\mathcal{U}} \mathcal{P} \quad &\Leftrightarrow \quad \forall U \in \mathcal{U} \quad \lim_{n\to\infty} \mu^U = \mu^U \quad (\text{in } \mathcal{P}(U)) \\ &\Leftrightarrow \quad \forall A \in \mathcal{E} \ \lim_{n\to\infty} h(\mu)(A) = h(\mu)(A)\end{aligned}$$

This ends the proof of the first part of the lemma. The proof of the second part can be found in Groeneboom, Oosterhoff and Ruymgaart (1979), and in Pinsker (1964)). ∎

The next technical lemma provides large deviations probabilities for independent and identically distributed sequences, on finite state space models.

LEMMA 1.7 *Let S be a finite state space, and let $Y = (Y^i)_{1\leq i\leq N}$ be a collection of N independent, and S-valued random variables, identically distributed according to a measure $\eta \in \mathcal{P}(S)$. For any $\mu \in m_N(S^N)$, we have*

$$(N+1)^{-|S|} \ \leq \ \exp\{N\mathrm{Ent}(\mu|\eta)\} \ \mathbb{P}_\eta(m_N(Y) = \mu) \ \leq 1$$

Proof:
This result is rather well known, see for instance Lemma 2.1.9 p. 15 in Dembo and Zeitouni (1993). Its proof is rather elementary. We notice that for $y \in m_N^{-1}(\mu)$ with $\mu \in m_N(S^N)$, we have

$$\mathbb{P}_\eta(Y = y) = \prod_{u\in S} \eta(u)^{N\mu(u)} = \exp N \sum_{u\in S} \mu(u) \log \eta(u)$$

This yields that

$$\begin{aligned}\mathbb{P}_\eta(m_N(Y) = \mu) \quad &= \quad \mathbb{P}_\eta(Y \in m_N^{-1}(\mu)) \\ &= \quad |m_N^{-1}(\mu)| \ \exp N \sum_{u\in S} \mu(u) \log \eta(u)\end{aligned} \tag{1.9}$$

Using the fact that $\mathbb{P}_\mu(m_N(Y) = \mu) \leq 1$, we find that

$$|m_N^{-1}(\mu)| \leq \exp -N \sum_{u\in S} \mu(u) \log \mu(u)$$

from which we conclude that $\mathbb{P}_\eta(m_N(Y) = \mu) \leq \exp\{-N\mathrm{Ent}(\mu|\eta)\}$. Since $n!/p! \geq p^{n-p}$, for any $n, p \geq 0$, one checks that for any $\eta \ll \mu$

$$\frac{\mathbb{P}_\mu(m_N(Y) = \mu)}{\mathbb{P}_\mu(m_N(Y) = \eta)} \quad = \quad [\prod_{u\in S} \frac{(N\eta(u))!}{(N\mu(u))!}] \ \mu(u)^{N(\mu(u)-\eta(u))} \geq 1$$

Recalling that $\mathrm{Ent}(\mu|\eta) \geq 0$, this also yields that

$$\mathbb{P}_\mu(m_N(Y) = \eta) \vee \mathbb{P}_\eta(m_N(Y) = \mu) \leq \mathbb{P}_\mu(m_N(Y) = \mu)$$

This implies that

$$1 = \mathbb{P}_\mu(m_N(Y) \in m_N(S^N)) \leq |m_N(S^N)| \mathbb{P}_\mu(m_N(Y) = \mu)$$

Using the fact that any $m_N(y) \in m_N(S^N)$ can be rewritten as

$$m_N(y) = \frac{1}{N} \sum_{u \in S} |\{i \ : \ y^i = u\}| \ \delta_u$$

we find that $|m_N(S^N)| \leq (N+1)^{|S|}$, from which we conclude that

$$|m_N^{-1}(\mu)| \geq (N+1)^{-|S|} \ \exp -N \sum_{u \in S} \mu(u) \log \mu(u)$$

This, together with (1.9), ends the proof of the lemma. ∎

THEOREM 1.8 *Let $(X^i)_{i \geq 1}$ be a sequence of independent, E-valued random variables, identically distributed according to a measure $\eta \in \mathcal{P}(E)$. For any N, we denote by Q^N the law on $\mathcal{P}(E)$ of the N-empirical measures $m_N(X)$. For any $U \in \mathcal{U}$, the sequence of distributions $Q^N \circ q_U^{-1}$ satisfy the large deviations principles on $(\mathcal{P}(U), d_U)$, with the good entropy rate function*

$$I_U = \mathrm{Ent}_U(\,.\,|\eta)$$

Combining Corollary 1.5 with Lemma 1.6, and recalling that the relative entropy is a good rate function for the τ-topology on $\mathcal{P}(E)$ (see for instance Groeneboom, Oosterhoff and Ruymgaart (1979) or p. 240-243 in Dembo and Zeitouni (1993)), we easily obtain the strong version of Sanov's Theorem, which states that Q^N satisfy the large deviations principles on $\mathcal{P}(E)$, equipped with the τ-topology, and with the good entropy rate function $\mathrm{Ent}(.|\eta)$.

In the reverse angle, by a simple contraction argument, Sanov's Theorem implies Corollary 1.5. To check that $I_U \geq \mathrm{Ent}_U(\nu|\eta)$, we use the fact that for any $\nu \in \mathcal{P}(U)$ we have

$$\mathrm{Ent}(\Omega_\eta(\nu)|\eta) = \mathrm{Ent}_U(\nu|\eta)$$

where Ω_η is the mapping from $\mathcal{P}(U)$ into $q_U^{-1}(\mathcal{P}(U))$, which associate with any $\nu \in \mathcal{P}(U)$, the measure $\Omega_\eta(\nu)$ defined for any $A \in \mathcal{E}$ and $i \leq d$ by

$$\Omega_\eta(\nu)(A \cap U^i) = \begin{cases} \eta(A \cap U^i) \ \nu(U^i)/\eta(U^i) & \text{if} \quad \eta(U^i) > 0 \\ \nu(A \cap U^i) & \text{if} \quad \eta(U^i) = 0 \end{cases} \tag{1.10}$$

Proof of Theorem 1.8:
Let $\Lambda_U : \mathcal{B}_b(U) \to \mathbb{R}$ be the logarithmic moment generating function defined for any $f \in \mathcal{B}_b(U)$ by

$$\begin{aligned}
\Lambda_U(f) &= \frac{1}{N} \log \int e^{N <f,\mu>} (Q^N \circ q_U^{-1})(d\mu) \\
&= \frac{1}{N} \log \mathbb{E}_\eta^N (e^{N m_N(X)(f)}) \\
&= \log \eta(\exp f) = \log q_U(\eta)(\exp f) < \infty
\end{aligned}$$

By Cramér's Theorem, the deviation upper bound follows from the fact that $\mathcal{P}(U)$ is compact, and $\mathrm{Ent}_U(\,.\,|\eta)$ is the Fenchel-Legendre transform of the function Λ_U.

To prove the lower bound, we let $A \subset \mathcal{P}(U)$ be a open set such that $\mathrm{Ent}_U(A|\eta)$ is finite (otherwise the proof of the lower bound is trivial). For any $\delta > 0$, there exists a point $\mu \in A$ such that

$$\mathrm{Ent}_U(\mu|\eta) \leq \mathrm{Ent}_U(A|\eta) + \delta \tag{1.11}$$

Since $\mu \in A$, and A is open in $(\mathcal{P}(U), d_U)$, there exists some $\varepsilon > 0$, such that

$$\mathcal{V}_U(\mu, \varepsilon) = \{\nu \in \mathcal{P}(U) \ : \ d_U(\nu, \mu) < \varepsilon\} \subset A$$

Up to a change of index, we suppose that $\mu(U^d) = \vee_{i=1}^d \mu(U^i) (> 0)$ and we associate with μ, the N-approximating distributions $\mu^N \in \mathcal{P}(U)$ defined by

$$\mu^N(U^i) = \begin{cases} [N\mu(U^i)]/N & \text{if} \ \ 1 \leq i < d \\ 1 - \sum_{i=1}^{d-1} [N\mu(U^i)]/N & \text{if} \ \ i = d \end{cases}$$

By construction, we note that

$$\mu^N \ll \mu \ (\ll q_U(\eta)) \quad \text{and} \quad \mathrm{Ent}_U(\mu^N|\eta) < \infty$$

Since $d_U(\mu^N, \mu) \leq (d-1)/N$, then we have $\mu^N \in \mathcal{V}_U(\mu, \varepsilon)$ for any $N > N_1 = (d-1)/\varepsilon$. On the other hand, we have

$$q_U^{-1}(\{\eta \in \mathcal{P}(U) \ : \ d_U(\eta, \mu^N) = 0\}) \ \subset \ q_U^{-1}(\mathcal{V}_U(\mu, \varepsilon))$$

and by Lemma 1.7, we get

$$\begin{aligned}
Q^N(q_U^{-1}(\mathcal{V}_U(\mu, \varepsilon))) &\geq \mathbb{P}_\eta^N(d_U(m_N(X), \mu^N) = 0) \\
&= \frac{N!}{(N\mu^N(U^1))! \dots (N\mu^N(U^d))!} \prod_{i=1}^d \eta(U^i)^{(N\mu^N(U^i))} \\
&\geq (N+1)^{-d} \ \exp\{-N \mathrm{Ent}_U(\mu^N|\eta)\}
\end{aligned}$$

The continuity of the entropy function $\mathrm{Ent}_U(.|\eta)$ on the set of measures $\{\mu \in \mathcal{P}(U) \ : \ \mu \ll q_U(\eta)\}$, now implies that for any $N \geq N_2$, and some $N_2 \geq 1$,

$$\mathrm{Ent}_U(\mu^N \mid \eta) \leq \mathrm{Ent}_U(\mu \mid \eta) + \delta$$

We finally conclude that for any $N \geq (N_1 \vee N_2)$

$$\begin{aligned} \frac{1}{N} \log Q^N(q_U^{-1}(A)) &\geq \frac{1}{N} \log \mathbb{Q}^N(q_U^{-1}(\mathcal{V}_U(\mu, \varepsilon))) \\ &\geq -\mathrm{Ent}_U(A|\eta) - d\, \frac{\log (N+1)}{N} - 2\delta \end{aligned}$$

Letting $N \to \infty$, and then $\delta \to 0$, the end of the proof of the lower bound is completed. This ends the proof of Theorem 1.8. ■

5. Large deviations for interacting processes

5.1 Some topological preliminaries

We equip the set $\mathcal{U}^n = \prod_{p=0}^n \mathcal{U}_p$ with the partial ordering defined by

$$(U_p)_{p\leq n} \leq (V_p)_{p\leq n} \Longleftrightarrow (U_p \leq V_p, \quad \forall 0 \leq p \leq n)$$

We associate with each $U^n = (U_p)_{p\leq n} \in \mathcal{U}^n$, the Cartesian product $\mathcal{P}^n(U) = \prod_{p=0}^n \mathcal{P}(U_p)$. For any $U^n \geq V^n$, we denote respectively by q_{U^n}, and q_{U^n,V^n}, the continuous and canonical projections from $\mathbf{P}^n(E)$, and respectively from $\mathcal{P}^n(U)$, into $\mathcal{P}^n(V)$. From the projective limit Theorem 1.6, it is easily established that the Cartesian product space $\mathbf{P}^n(E)$, equipped with the product τ-topology, is homeomorphic to the projective limit $\lim_{\mathcal{U}^n} \mathcal{P}^n$ of the set

$$\mathcal{P}^n = (\mathcal{P}^n(U), q_{U^n,V^n})_{U^n \geq V^n}$$

For any $n \geq 0$, we denote by J_n, the function on $\mathbf{P}^n(E) = \prod_{p=0}^n \mathbf{P}(E_p)$ defined by

$$J_n((\mu_p)_{0\leq p\leq n}) = H(\mu_0|\eta_0) + \sum_{p=1}^n H(\mu_p|\Phi_p(\mu_{p-1}))$$

with the convention $\Phi_0(\mu_{-1}) = \eta_0$ for $p = 0$. Observe that J_n is lower semi-continuous, since it can be obtained as the supremum of continuous functions

$$\begin{aligned} &J_n((\mu_p)_{0\leq p\leq n}) \\ &= \sup\{\textstyle\sum_{p=0}^n [\mu_p(f_p) - \log \Phi_p(\mu_{p-1})(e^{f_p})] \quad (f_p)_{p\leq n} \in \cup_{U^n \in \mathcal{U}^n} \mathcal{B}_b^n(U)\} \end{aligned}$$

where $\mathcal{B}_b^n(U) = \prod_{p=0}^n \mathcal{B}_b(U_p)$.

On the other hand, by Theorem 1.6, and using a simple induction on the time parameter, we prove that the domain $\{\mu; J_n(\mu) < \infty\}$ of J_n is included in $\mathcal{P}^n(E)$. In addition for any $(\mu_p)_{p\leq n} \in \mathcal{P}^n(E)$ we have

$$J_n((\mu_p)_{0\leq p\leq n}) = \mathrm{Ent}(\mu_0|\eta_0) + \sum_{p=1}^n \mathrm{Ent}(\mu_p|\Phi_p(\mu_{p-1}))$$

This implies that for any τ_1-closed, or τ_1-open subset, A of $\mathbf{P}^n(E)$, we have $J_n(A \cap \mathcal{P}^n(E)) = J_n(A)$.

The above observations, combined with Theorem 1.1, imply that Q_n^N satisfies the large deviations principles on $\mathcal{P}^n(E)$, equipped with the relative τ-topology, with rate function J_n. To prove that J_n is a good rate function, we use the lower semi-continuity property of J_n on $\mathbf{P}^n(E)$ to check that the level sets $J_n^{-1}([0,a])$, $a > 0$, are τ_1-closed, thus τ_1-compact in the compact set $\mathbf{P}^n(E)$. Since we proved that $J_n^{-1}([0,a]) \subset \mathcal{P}^n(E)$, we conclude that these level sets are also τ-compact subsets of $\mathcal{P}^n(E)$.

It is also convenient to notice that Corollary 1.5, combined with Theorem 1.1, implies that for any $U^n = (U_p)_{p\leq n} \in \mathcal{U}^n$, the image measures $Q_n^N \circ q_{U^n}^{-1}$ satisfy the large deviations principles on $\mathcal{P}^n(U)$, with the good rate function defined for any $\nu = (\nu_p)_{p\leq n} \in \mathcal{P}^n(U)$ by

$$J_n^U(\nu) = J_n(q_{U^n}^{-1}(\nu)) \ (= \inf_{q_{U^n}^{-1}(\nu)} J_n)$$

Furthermore, arguing as in (1.10), we find that

$$J_n^U(\nu) = \sum_{p=0}^n \inf_{\mu_{p-1}\in q_{U_{p-1}}^{-1}(\nu_{p-1})} \mathrm{Ent}_{U_p}(\nu_p|\Phi_p(\mu_{p-1}))$$

with the conventions $q_{U_{-1}}^{-1}(\nu_{-1}) = \{\eta_0\}$, and $\Phi_0 = Id$ for $p = 0$.

We end this section with an elementary proof of Corollary 1.2. We first observe that for any flow $\mu = (\mu_p)_{p\leq n} \in \mathcal{P}^n(E)$, we have

$$d_{U_n}(\mu_n, \eta_n) \leq \|\mu_n - \nu_n\|_{tv}$$

Using the triangular inequality, and condition (L), we find that

$$\begin{aligned}
\|\mu_n - \eta_n\|_{tv} &\leq \sum_{p=0}^n \|\Phi_{p,n}(\mu_p) - \Phi_{p,n}(\Phi_p(\mu_{p-1}))\|_{tv} \\
&\leq \sum_{p=0}^n \beta(\Phi_{p,n}) \ \|\mu_p - \Phi_p(\mu_{p-1})\|_{tv}
\end{aligned}$$

with the convention $\Phi_0(\mu_{-1}) = \eta_0$ for $p = 0$. This readily yields that

$$d_{U_n}(\mu_n, \eta_n)^2 \leq \beta(\Phi) \sum_{p=0}^{n} \|\mu_p - \Phi_p(\mu_{p-1})\|_{tv}^2 \leq \beta(\Phi)\ J_n(\mu)$$

In the last inequality, we have used the fact that the relative entropy dominates the square of the total variation distance (see for instance Csiszar (1967), Theorem 4.1). From the additive structure of J_n, we conclude that

$$\sup_{p\leq n} d_{U_p}(\mu_p, \eta_p)^2 \leq \beta(\Phi)\ J_n(\mu)$$

The end of the proof of the cororally is now clear.

5.2 Proof of Theorem 1.1

To prepare the proof of the upper bound, we again recall that any closed F of $\mathbf{P}^n(E)$ has the form $F = \cap_{U^n \in \mathcal{U}^n} q_{U^n}^{-1}(F_{U^n})$, where F_{U^n} stands for the τ_1-closure of the set $q_{U^n}(F)$. Since $\mathbf{P}^n(E)$ is compact, the sets $q_{U^n}^{-1}(F_{U^n})$ are compact, and for any $\varepsilon > 0$, one can find a finite and open ε-covering

$$q_{U^n}^{-1}(F_{U^n}) \subset \cup_{i=1}^{m} \mathcal{V}_n^{U^n}(\mu^i, \varepsilon)$$

with $\mu^i = (\mu_p^i)_{p\leq n} \in F$ and $\mathcal{V}_p^{U^n}(\mu^i, \varepsilon) = (\prod_{0\leq q\leq p} \mathcal{V}_p^{U^n}(\mu_q^i, \varepsilon))$, for any $p \leq n$. Under our continuity assumptions, we can choose these open neighborhoods such that for any $\eta_p \in \mathcal{V}_p^{U^n}(\mu_p^i, \varepsilon)$.

$$d_{U_p}(\eta_p, \mu_p^i) \vee d_{U_{p+1}}(\Phi_{p+1}(\eta_p), \Phi_{p+1}(\mu_p^i)) \leq \varepsilon$$

In this case, we have for each $(f_p)_{p\leq n} \in \mathcal{B}_b^n(U_n)$

$$\begin{aligned}
&\mathbb{P}((\eta_p^N)_{p\leq n} \in \mathcal{V}_n^{U^n}(\mu^i, \varepsilon)) \\
&= \mathbb{E}(e^{N(\eta_n^N - \mu_n^i)(f_n) - N(\eta_n^N - \mu_n^i)(f_n)}\ 1_{\mathcal{V}_{n-1}^{U_n}(\mu^i,\varepsilon)}((\eta_p^N)_{p<n})) \\
&\leq e^{\varepsilon N\|f_n\|}\ \mathbb{E}(e^{N(\eta_n^N - \mu_n^i)(f_n)}\ 1_{\mathcal{V}_{n-1}^{U_n}(\mu^i,\varepsilon)}((\eta_p^N)_{p<n}))
\end{aligned}$$

By definition of the particle approximating model, we arrive at

$$\begin{aligned}
&\mathbb{P}((\eta_p^N)_{p\leq n} \in \mathcal{V}_n^{U^n}(\mu^i, \varepsilon)) \\
&\leq e^{\varepsilon N\|f_n\|}\ \mathbb{E}(e^{-N(\mu_n^i(f_n) - \log\Phi_n(\eta_{n-1}^N)(e^{f_n}))}\ 1_{\mathcal{V}_{n-1}^{U_n}(\mu^i,\varepsilon)}((\eta_p^N)_{p<n})) \\
&\leq e^{\varepsilon N c(f_n) - N(\mu_n^i(f_n) - \log\Phi_n(\mu_{n-1}^i)(e^{f_n}))}\ \mathbb{P}((\eta_p^N)_{p<n} \in \mathcal{V}_{n-1}^{U_n}(\mu^i, \varepsilon))
\end{aligned}$$

for some finite constant $c(f_n) < \infty$, whose values only depend on the supremum norm of f_n. A simple induction now yields that

$$\frac{1}{N}\log \mathbb{P}((\eta_p^N)_{p\le n} \in \mathcal{V}_n^{U^n}(\mu^i,\varepsilon))$$

$$\le \varepsilon \textstyle\sum_{p\le n} c(f_p) - \sum_{p\le n}(\mu_p^i(f_p) - \log \Phi_p(\mu_{p-1}^i)(e^{f_p}))$$

from which we conclude that

$$\limsup_{\varepsilon\to 0}\limsup_{N\to\infty} \frac{1}{N}\log \mathbb{P}((\eta_p^N)_{p\le n} \in \mathcal{V}_n^{U^n}(\mu^i,\varepsilon))$$

$$\le -\textstyle\sum_{p\le n}(\mu_p^i(f_p) - \log \Phi_p(\mu_{p-1}^i)(e^{f_p}))$$

Taking the infimum over all $(f_p)_{p\le n} \in \mathcal{B}_b^n(U_n)$ in the r.h.s of the above display, we conclude that

$$\limsup_{\varepsilon\to 0}\limsup_{N\to\infty} \frac{1}{N}\log \mathbb{P}((\eta_p^N)_{p\le n} \in \mathcal{V}_n^{U^n}(\mu^i,\varepsilon))$$

$$\le -\textstyle\sum_{p\le n} \mathrm{Ent}_{U_n}(\mu_p^i | \log \Phi_p(\mu_{p-1}^i))$$

$$\le -J_n^U(q_{U^n}(\mu^i)) \le -J_n^U(F_{U^n})$$

By the union of events bounds, we find that

$$\limsup_{N\to\infty} \frac{1}{N}\log \mathbb{P}((\eta_p^N)_{p\le n} \in q_{U^n}^{-1}(F_{U^n})) \le -J_n^U(F_{U^n})$$

and therefore

$$\limsup_{N\to\infty} \frac{1}{N}\log \mathbb{P}((\eta_p^N)_{p\le n} \in F) \le -J_n^U(F_{U^n}) \le -J_n^U(F_{U^n})$$

We end the proof of the upper bound by taking the infimum over all U^n, and invoquing the min-max Theorem 1.4.

Our final objective is to prove that the rate function J_n governs the large deviations principles lower bounds. To this end, we let $A \subset \mathbf{P}^n(E)$ be a τ_1-open set such that $J_n(A) = \inf_A J_n < \infty$ (otherwise the proof of the lower bound is as usually trivial). For any $\delta > 0$, there exists a point $\mu \in A \cap \mathcal{P}^n(E)$, such that

$$J_n(\mu) \le J_n(A) + \delta \tag{1.12}$$

Since $\mu = (\mu_p)_{0\le p\le n} \in A$, we can find a collection of strictly positive numbers $(\varepsilon_p)_{0\le p\le n}$, and a sequence of finite partitions $U_p = (U_p^i)_{1\le i\le d_p}$ of the sets E_p, such that

$$\mathcal{C}_{\varepsilon,n}^U(\mu) = \prod_{0\le p\le n} B_{U_p}(\mu_p,\varepsilon_p) \subset A$$

with the open neighborhood $B_{U_p}(\mu_p, \varepsilon_p)$ of $\mu_p \in \mathbf{P}(E_p)$ given by

$$B_{U_p}(\mu_p, \varepsilon_p) = \{\nu_p \in \mathbf{P}(E_p) \ : \ d_{U_p}(\mu_p, \nu_p) < \varepsilon_p\}$$

Up to a change of index, we suppose that $\mu_n(U_n^d) = \vee_{i=1}^{d_n} \mu_n(U_n^i)(> 0)$, and we associate with $\mu_n \in \mathbf{P}(E_n)$, the N-approximating distributions $\mu_n^N \in \mathcal{P}(U_n)$ defined by

$$\mu_n^N(U_n^i) = \begin{cases} [N\mu_n(U_n^i)]/N & \text{if} \quad 1 \leq i < d_n \\ 1 - \sum_{i=1}^{d_n-1} [N\mu_n(U_n^i)]/N & \text{if} \quad i = d_n \end{cases}$$

By construction, we note that

$$\mu_n^N \ll q_{U_n}(\mu_n) \ll q_{U_n}(\Phi_n(\mu_{n-1})) \quad \text{and} \quad \mathrm{Ent}_{U_n}(\mu_n^N | \Phi_n(\mu_{n-1})) < \infty$$

Under our assumptions, we also have for any $\nu_{n-1} \in B_{U_{n-1}}(\mu_{n-1}, \varepsilon_{n-1})$, and $1 \leq i \leq d_n$,

$$\mu_n^N(U_n^i) > 0 \Longrightarrow \Phi_n(\nu_{n-1})(U_n^i) > 0$$

Since $d_{U_n}(\mu_n, \mu_n^N) \leq (d_n - 1)/N$, we find that $\mu_n^N \in B_{U_n}(\mu_n, \varepsilon_n)$ as soon as $N > N_n^1 = (d_n - 1)/\varepsilon_n$, and hence

$$\mathcal{C}_{\varepsilon,n-1}^U((\mu_k)_{k<n}) \times \{\nu_n \in \mathbf{P}(E_n) \ : \ d_{U_n}(\mu_n^N, \nu_n) = 0\} \subset A$$

By definition of the N-particle model, and using Lemma 1.7, it is also clear that

$$\begin{aligned} &\mathbb{P}(d_{U_n}(\mu_n^N, \eta_n^N) = 0 \mid \eta_{n-1}^N) \\ &= \frac{N!}{(N\mu_n^N(U_n^1))! \ldots (N\mu_n^N(U_n^{d_n}))!} \prod_{i=1}^{d_n} \Phi_n(\eta_{n-1}^N)(U_n^i)^{(N\mu_n^N(U_n^i))} \\ &\geq \exp\left[-N \ \mathrm{Ent}_{U_n}(\mu_n^N \mid \Phi_n(\eta_{n-1}^N)) - d \log{(N+1)}\right] \end{aligned}$$

On the other hand, we have for any $\nu_{n-1} \in B_{U_{n-1}}(\mu_{n-1}, \varepsilon_{n-1}) \cap \mathcal{P}(E_{n-1})$ the inequality

$$\begin{aligned} &\mathrm{Ent}_{U_n}(\mu_n^N \mid \Phi_n(\nu_{n-1})) \\ &\leq \mathrm{Ent}_{U_n}(\mu_n^N \mid \Phi_n(\mu_{n-1})) + \textstyle\sum_{i:\mu_n^N(U_n^i)>0} \mu_n^N(U_n^i) \log \frac{d\Phi_n(\mu_{n-1})(U_n^i)}{d\Phi_n(\nu_{n-1})(U_n^i)} \end{aligned}$$

Under our assumptions, we have

$$\begin{aligned} &\mathrm{Ent}_{U_n}(\mu_n^N \mid \Phi_n(\nu_{n-1})) \\ &\leq \mathrm{Ent}_{U_n}(\mu_n^N \mid \Phi_n(\mu_{n-1})) + \tfrac{2}{\rho_n \ \lambda_n^\star(U_n)} \ \ d_{U_n}(\Phi_n(\nu_{n-1}), \Phi_n(\mu_{n-1})) \end{aligned}$$

with $\lambda_n^\star(U_n) = \wedge_{i:\lambda_n(U_n^i)>0}\lambda_n(U_n^i)$. Since the mapping Φ_n is τ_1-continuous, for every $\delta > 0$, there exists some τ_1-open neighborhood $\mathcal{O}_\delta(\mu_{n-1}) \subset \mathbf{P}(E_{n-1})$ of μ_{n-1} (which may depend on U_n), such that

$$\mathcal{O}_\delta(\mu_{n-1}) \subset B_{U_{n-1}}(\mu_{n-1}, \varepsilon_{n-1})$$

and on the set of events $\{\eta_{n-1}^N \in \mathcal{O}_\delta(\mu_{n-1})\}$, we have

$$\mathrm{Ent}_{U_n}(\mu_n^N \mid \Phi_n(\eta_{n-1}^N)) \quad \leq \quad \mathrm{Ent}_{U_n}(\mu_n^N \mid \Phi_n(\mu_{n-1})) + \delta \ (< \infty)$$

The continuity of the function $\mathrm{Ent}_{U_n}(.\mid \Phi_n(\mu_{n-1}))$, on the set of measures $\{\nu_n \in \mathcal{P}(U_n) \ : \ \nu_n \ll q_{U_n}(\Phi_n(\mu_{n-1}))\}$, now implies that

$$\begin{aligned} \mathrm{Ent}_{U_n}(\mu_n^N \mid \Phi_n(\mu_{n-1})) &\leq \mathrm{Ent}_{U_n}(\mu_n \mid \Phi_n(\mu_{n-1})) + \delta \\ &\leq H(\mu_n \mid \Phi_n(\mu_{n-1})) + \delta \end{aligned}$$

for any $N \geq N_n^2$, and some $N_n^2 \geq 1$. It is now convenient to observe that

$$\begin{aligned} Q_n^N(A) &= \mathbb{P}_{\eta_0}^N((\eta_0^N, \ldots, \eta_n^N) \in A) \\ &\geq \mathbb{P}_{\eta_0}^N \big(d_{U_n}(\mu_n^N, \eta_n^N) = 0 \ , \ \eta_{n-1}^N \in \mathcal{O}_\delta(\mu_{n-1}) \\ &\qquad\qquad \text{and} \ \ (\eta_k^N)_{p<n-1} \in \mathcal{C}_{\varepsilon,n-2}^U((\mu_p)_{p\leq n-2}\big) \end{aligned}$$

Using a simple calculation, we deduce the following lower bound

$$\begin{aligned} &Q_n^N(A) \\ &\geq \inf_{\nu\in\mathcal{O}_\delta(\mu_{n-1})} \mathbb{P}\left(d_{U_n}(\mu_n^N, \eta_n^N) = 0 \mid \eta_{n-1}^N = \nu\right) \quad Q_{n-1}^N\left(\mathcal{D}_{(\delta,\varepsilon),n-1}(\mu)\right) \end{aligned}$$

with the τ_1-open neighborhood $\mathcal{D}_{(\delta,\varepsilon),n-1}(\mu) \in \mathbf{P}^{n-1}(E)$ of $(\mu_p)_{p\leq n-1}$ defined by

$$\mathcal{D}_{(\delta,\varepsilon),n-1}(\mu) = \mathcal{C}_{\varepsilon,n-2}^U((\mu_p)_{p\leq n-2}) \times \mathcal{O}_\delta(\mu_{n-1})$$

¿From previous estimations, we find that, for any $N \geq N_n^1 \vee N_n^2$

$$\begin{aligned} &\tfrac{1}{N}\log Q_n^N(A) \\ &\geq -\ H(\mu_n \mid \Phi_n(\mu_{n-1})) - \tfrac{d\log(N+1)}{N} - 2\delta + \tfrac{1}{N}\log Q_{n-1}^N\left(\mathcal{D}_{(\delta,\varepsilon),n-1}(\mu)\right) \\ &= -J_n(\mu) - \tfrac{d\log(N+1)}{N} - 2\delta + J_{n-1}((\mu_p)_{p\leq n}) + \tfrac{1}{N}\log Q_{n-1}^N\left(\mathcal{D}_{(\delta,\varepsilon),n-1}(\mu)\right) \end{aligned}$$

Furthermore by (1.12), and using our induction hypothesis, we have

$$\begin{aligned} &\liminf_{N\to\infty} \tfrac{1}{N}\log Q_n^N(A) \\ &\geq -J_n(A) - 3\delta + [J_{n-1}((\mu)_{p<n}) - J_{n-1}(\mathcal{D}_{(\delta,\varepsilon),n-1}(\mu))] \geq -J_n(A) - 3\delta \end{aligned}$$

Letting δ tends to 0, we conclude that J_n governs the lower bounds. ■

References

I. Csiszar. Information type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar*, vol. 2, pp. 299-318 (1967).

D. A.Dawson and J. Gärtner. Analytic aspects of multilevel large deviations. *Asymptotic methods in probability and statistics* (Ottawa,ON,1997), pp. 401–440, North-Holland Amsterdam (1998).

D.A. Dawson and S. Feng. *Large deviations for the Fleming-Viot process with neutral mutation and selection, II,* Stochastic Processes and their Applications, vol. 92, pp. 131–162 (2001).

D.A. Dawson and J. Gärtner. *Large deviations from the McKean Vlasov limit for weakly interacting diffusions* Stochastics, pp. 247-308,vol. 20 (1987).

A. Dembo and O. Zeitouni. Large deviations techniques and Applications. *Jones and Bartlett* (1993).

A. de Acosta. Projective systems in large deviation theory. II. Some applications. *Probability in Banach spaces, 9 (Sandjberg, 1993)*, pp. 241–250, Progr. Probab., vol. 35, Birkhäuser Boston, (1994).

A. de Acosta. Exponential tightness and projective systems in large deviation theory. *Festschrift for Lucien Le Cam*, pp. 143-156, Springer, New York, (1997).

A. de Acosta. On large deviations of empirical measures in the τ-topology. *J. Appl. Probab.*, Studies in applied probability, vol. 31A, pp. 41-47 (1994).

P. Groeneboom, J. Oosterhoff and F.H. Ruymgaart. Large Deviation Theorems for Empirical Probability Measures. *The Annals of Proba.*, vol. 7, no 4, pp. 553–586 (1979).

P. Del Moral and A. Guionnet. Large deviations for interacting particle systems. Applications to non linear filtering problems. *Stochastic Processes and their Applications*, vol. 78, pp. 69-95 (1998).

P. Del Moral and L. Miclo, Branching and Interacting Particle Systems Approximations of Feynman-Kac Formulae with Applications to Non-Linear Filtering. *Séminaire de Probabilités XXXIV, Lecture Notes in Mathematics 1729, Springer-Verlag Berlin*, Ed. J. Azéma and M. Emery and M. Ledoux and M. Yor, pp. 1-145 (2000).

P. Del Moral and L. Miclo, Genealogies and Increasing Propagations of Chaos for Feynman-Kac and Genetic Models. *The Annals of Applied Probability*, vol. 11, No. 4, pp. 1166-1198 (2001).

P. Del Moral. Measure valued processes and interacting particle systems. Application to non linear filtering problems. *The Annals of Applied Probability*, vol. 8, no. 2, pp. 438–495 (1998).

P. Del Moral and T. Zajic. A note on the Laplace-Varadhan's integral Lemma. *Bernoulli*, vol. 9, no. 1, pp. 49–65 (2003).

Ed. A. Doucet and N. De Freitas and N. Gordon. Sequential Monte Carlo Methods in Pratice. *Springer, Statistics for engineering and Information Science* (2001).

S. Méléard, Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic Models for Nonlinear Partial Differential Equations* Editors D. Talay, and L. Tubaro. pp. 43–95, Springer-Verlag, Berlin-Heidelberg.

M.S. Pinsker. Information and information stability of random variables and processes, *Holden Day, San Francisco* (1964).

A.S. Sznitman. Topics in propagation of chaos. *École d'été de Probabilités de Saint-Flour XIX*, Lect. Notes in Math. no. 1464, Springer-Verlag, Berlin-Heidelberg (1991).

K.P. Körding and D.M. Wolpert, *Nature*, 427, 244, (2004).