Ann. I. H. Poincaré – PR 38, 5 (2002) 649–680

S0246-0203(01)01094-9/FLA

LARGE DEVIATIONS FOR A TRIANGULAR ARRAY OF EXCHANGEABLE RANDOM VARIABLES

GRANDES DÉVIATIONS POUR UN TABLEAU TRIANGULAIRE DE VARIABLES ALÉATOIRES ÉCHANGEABLES

José TRASHORRAS

Laboratoire de Probabilités et Modèles Aléatoires, Université Paris 7, UFR de Mathématiques, case 7012, 2 Place Jussieu, 75251 Paris Cedex 05, France

Received 20 September 2000, revised 3 April 2001

ABSTRACT. – In this paper we consider a triangular array whose rows are composed of finite exchangeable random variables. We prove that, under suitable conditions, the sequence defined by the empirical measure process of each row satisfies a large deviation principle. We first study the particular case where the rows are given by sampling without replacement from fixed urns. Then we prove a large deviation principle in the general setting, by identifying finite exchangeable random variables and sampling without replacement from urns with random composition. © 2002 Éditions scientifiques et médicales Elsevier SAS

AMS classification: Primary 60F10; secondary 60G09 and 62G09

Keywords: Large deviations; Exchangeable random variables; Sampling without replacement

RÉSUMÉ. – Nous considérons un tableau triangulaire dont les lignes sont composées de variables aléatoires fini-échangeables. Nous prouvons sous certaines conditions que la suite définie par le processus de mesure empirique de chaque ligne vérifie un principe de grandes déviations. Dans un premier temps nous traitons le cas particulier où chaque ligne résulte du tirage sans remise dans une urne de composition donnée. Nous en déduisons ensuite un principe de grandes déviations dans le cas général, en identifiant les variables aléatoires fini-échangeables avec le tirage sans remise dans des urnes de composition aléatoire. © 2002 Éditions scientifiques et médicales Elsevier SAS

E-mail address: Jose.Trashorras@gauss.math.jussieu.fr (J. Trashorras).

1. Introduction

We say that a sequence of Borel probability measures $(P^n)_{n\in\mathbb{N}}$ on a topological space obeys a Large Deviation Principle (hereafter abbreviated LDP) with rate function I and in the scale $(a_n)_{n\in\mathbb{N}}$ if $(a_n)_{n\in\mathbb{N}}$ is a real-valued sequence satisfying $a_n \to \infty$ and I is a non-negative, lower semicontinuous function such that

$$-\inf_{x\in A^o} I(x) \leqslant \liminf_{n\to\infty} \frac{1}{a_n} \log P^n(A) \leqslant \limsup_{n\to\infty} \frac{1}{a_n} \log P^n(A) \leqslant -\inf_{x\in \bar{A}} I(x)$$

for any measurable set A, whose interior is denoted by A^o and closure by $\bar{A}$. Unless explicitly stated otherwise, we will take $a_n = n$. If the level sets $\{x\colon\ I(x) \leqslant \alpha\}$ are compact for every $\alpha < \infty$, I is called a good rate function. With a slight abuse of language we say that a sequence of random variables obeys a LDP when the sequence of measures induced by these random variables obeys a LDP. For a background on the theory of large deviations, see Dembo and Zeitouni [6] and references therein.

In this paper, we are interested in the LD behavior of *finite exchangeable* random variables. The word *exchangeable* appears in the literature for both *infinite exchangeable* sequences of random variables, and *finite exchangeable* random vectors. A sequence of random variables $(X_1,\dots,X_n,\dots)$ defined on a probability space $(\Omega,\mathcal{A},\mathbb{P})$ is *infinite exchangeable* if and only if for every permutation τ on $\mathbb{N}$ such that $|\{i, \tau(i) \neq i\}| < \infty$ the following identity in distribution holds

$$(X_1,\dots,X_n,\dots) \stackrel{\mathcal{D}}{=} (X_{\tau(1)},\dots,X_{\tau(n)},\dots).$$

An n-tuple $(X_1,\dots,X_n)$ of random variables defined on the same probability space is *finite exchangeable* or *n-exchangeable* (to indicate the number of random variables) if and only if for all permutations σ on $\{1,\dots,n\}$ it satisfies the identity in distribution

$$(X_1,\dots,X_n) \stackrel{\mathcal{D}}{=} (X_{\sigma(1)},\dots,X_{\sigma(n)}).$$

Finite and infinite exchangeability are related since any n-tuple extracted from an infinite exchangeable sequence of random variables is n-exchangeable. While LD for infinite exchangeable sequences have been entirely studied by Dinwoodie and Zabell [9], much less is known in the more intricate case of finite exchangeable random variables. After introducing our setting, we shortly review below known facts about exchangeable random variables. We refer to Aldous [1] for a large survey on this topic.

Throughout the sequel (Σ, d) will denote a Polish space, and $M^+(\Sigma)$ [resp. $M^1(\Sigma)$] the space of Borel non-negative measures [resp. probability measures] on Σ. These spaces will always be equipped with the topology of weak convergence, and we shall denote convergence in this topology by $\mu^n \stackrel{w}{\to} \mu$. Let us recall that the dual-bounded-Lipschitz metric β on $M^+(\Sigma)$ is compatible with this topology (see Dembo and Zajic [4], Appendix A.1).

De Finetti's well-known theorem (see, for example, [12]) states that any Σ valued infinite exchangeable sequence of random variables $(X_1,\dots,X_n,\dots)$ defined on

$(\Omega, \mathcal{A}, \mathbb{P})$ is a mixture of independent and identically distributed sequences of random variables, i.e. for any Borel set A of Σ^n

$$\mathbb{P}\big((X_1, \ldots, X_n) \in A\big) = \int_\Theta P_\theta\big((X_1, \ldots, X_n) \in A\big)\gamma(\mathrm{d}\theta),$$

where γ is a probability measure on a closed subset Θ of $M^1(\Sigma)$, and for every $\theta \in \Theta$, P_θ is a probability measure defined on $(\Omega, \mathcal{A})$ such that $X_1, \ldots, X_n, \ldots$ are independent and identically distributed under P_θ. Using this result, Dinwoodie and Zabell [9] have shown that if Θ is compact, the distribution of $\frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ under $\mathbb{P}$ satisfies a LDP with good rate function

$$I(\nu) = \inf_{\theta\in\Theta} H(\nu \mid \pi_\theta),$$

where $\pi_\theta = P_\theta \circ X_1^{-1}$ and $H(\cdot \mid \cdot)$ stands for the usual relative entropy (see Dupuis and Ellis [10] for a nice account on relative entropy).

Nevertheless, de Finetti's theorem is not valid for finite exchangeable random variables, as can be seen in the following simple example that arises in sampling. Consider an urn with n labelled balls $(x_1, \ldots, x_n)$. The result $(X_1, \ldots, X_n)$ of n draws without replacement among $(x_1, \ldots, x_n)$ is an n-exchangeable random vector that cannot be represented as a mixture of independent and identically distributed random variables. In this special case, Dembo and Zeitouni [5] have showed that if $\frac{1}{n}\sum_{i=1}^n \delta_{x_i} \xrightarrow{w} \mu$ then, for fixed $t_0 \in]0, 1[$, the distribution of $\frac{1}{[nt_0]}\sum_{i=1}^{[nt_0]} \delta_{X_i}$ follows a LDP in the scale $[nt_0]$ and with good rate function

$$I(\nu, t_0, \mu) = \begin{cases} H(\nu \mid \mu) + \frac{(1-t_0)}{t_0} H\big(\frac{\mu - t_0\nu}{1-t_0} \mid \mu\big) & \text{if } \frac{\mu - t_0\nu}{1-t_0} \in M^1(\Sigma), \\ \infty & \text{otherwise}. \end{cases}$$

Another well-known fact is that a family of n-exchangeable random variables can be approximated by n independent and identically distributed random variables in the variation norm (see Diaconis and Freedman [8]). However, this property does not give any hint for the LDP.

Here we consider a *finite exchangeable triangular array* $((X_i^n)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ of Σ valued random variables defined on $(\Omega, \mathcal{A}, \mathbb{P})$, i.e., *each row* $(X_1^n, \ldots, X_n^n)$ *is finite exchangeable*. We define the associated sequence of empirical measure processes by

$$L_t^n = \frac{1}{n}\sum_{i=1}^{[nt]} \delta_{X_i^n} \tag{1}$$

for every $t \in [0, 1]$. The process $(L_t^n)_{t\in[0,1]}$ belongs to the space $D[[0, 1], (M^+(\Sigma), \beta)]$ of all maps defined on $[0, 1]$ that are continuous from the right and have left limits. This space is endowed with the topology defined by the uniform metric

$$\beta_\infty(y_., z_.) = \sup_{t\in[0,1]} \beta(y_t, z_t), \tag{2}$$

where $y_.$ is a shortcut for $(y_t)_{t\in[0,1]}$.

The experience we are interested in can be heuristically described this way: From any n-tuple $(Y_i^n)_{1\leqslant i\leqslant n}$ of random variables one can simply obtain an n-exchangeable random vector $(X_i^n)_{1\leqslant i\leqslant n}$ by sampling without replacement from an urn with n labelled balls $(Y_1^n,\dots,Y_n^n)$. Equivalently, we let in this case $X_i^n = Y_{\sigma(i)}^n$, for $i=1,\dots,n$, with $\sigma=\sigma^n$ a random permutation on $\{1,\dots,n\}$ which is independent from $(Y_i^n)_{1\leqslant i\leqslant n}$ and uniformly distributed. Our purpose in this paper is to derive the LDP for $(L_t^n)_{t\in[0,1]}$ from the LDP for $\frac{1}{n}\sum_{i=1}^n \delta_{Y_i^n}$. Now, let us describe our setting rigorously. Let $\mathcal{B}_{\Sigma^n}$ be the Borel σ-algebra on Σ^n and P^n be any probability measure on $(\Sigma^n,\mathcal{B}_{\Sigma^n})$. We denote by $(Y_1^n,\dots,Y_n^n)$ the coordinate maps on $(\Sigma^n,\mathcal{B}_{\Sigma^n})$ when we consider them distributed according to P^n. Let $\mathbb{P}^n$ be the probability measure defined on every product $A_1\times\cdots\times A_n$ of measurable subsets of Σ by

$$\mathbb{P}^n(A_1\times\cdots\times A_n)=\frac{1}{n!}\sum_{\sigma\in S_n} P^n(A_{\sigma(1)}\times\cdots\times A_{\sigma(n)}), \tag{3}$$

where S_n is the symmetric group of order n. We denote by $(X_1^n,\dots,X_n^n)$ the coordinate maps on $(\Sigma^n,\mathcal{B}_{\Sigma^n})$ when its joint law is $\mathbb{P}^n$. Clearly, the random variables $(X_i^n)_{1\leqslant i\leqslant n}$ are n-exchangeable. Let $(\Omega,\mathcal{A},\mathbb{P})$ be the probability space associated to the sequence $((\Sigma^n,\mathcal{B}_{\Sigma^n},\mathbb{P}^n))_{n\in\mathbb{N}}$. Note that the mapping from Σ^n to $D[[0,1],(M^+(\Sigma),\beta)]$ defined by $(L_t^n)_{t\in[0,1]}$ is continuous, hence Borel measurable. As mentioned before, our goal is to derive the LDP [resp. the weak law of large numbers] for the distribution of $(L_t^n)_{t\in[0,1]}$ under $\mathbb{P}^n$ from the LDP [resp. the weak law of large numbers] for the distribution of $\frac{1}{n}\sum_{i=1}^n\delta_{Y_i^n}$ under P^n. Remark that [9] does not apply in this case.

The key to the proof is the following elementary fact. The law of $(X_1^n,\dots,X_n^n)$ conditioned on $\{\frac{1}{n}\sum_{i=1}^n\delta_{X_i^n}=\rho\}$, where ρ is an atomic measure whose atoms weigh $\frac{k}{n}$ $(1\leqslant k\leqslant n)$, is the law of sampling without replacement among these atoms counted with their frequency of appearance in ρ. Hence our analysis essentially reduces to the following particular case. Let $((y_i^n)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ be a fixed triangular array of elements of Σ, whose composition is given by $(\mu^n=\frac{1}{n}\sum_{i=1}^n\delta_{y_i^n})_{n\in\mathbb{N}}$, possibly with ties. For every $n\in\mathbb{N}$, we sample without replacement from the urn containing $(y_i^n)_{1\leqslant i\leqslant n}$ and we denote by x_i^n the ith element drawn. We call $\mathbb{P}^n(\cdot\,;\mu^n)$ the distribution on Σ^n related to this sampling. For every $n\in\mathbb{N}$ it clearly makes $(x_i^n)_{1\leqslant i\leqslant n}$ a finite exchangeable vector. For all $t\in[0,1]$ we set

$$l_t^n=\frac{1}{n}\sum_{i=1}^{[nt]}\delta_{x_i^n}, \tag{4}$$

and for all $\mu\in M^1(\Sigma)$ we let $\mathcal{AC}_\mu$ be the space of all maps $\nu_t:[0,1]\to M^+(\Sigma)$ such that:

1. $\nu_t-\nu_s\in M^+(\Sigma)$ is of total mass $t-s$ for all $0\leqslant s\leqslant t\leqslant 1$.
2. $\nu_0=0$ and $\nu_1=\mu$.
3. $\nu_\cdot$ possesses a *weak derivative* for almost every $t\in[0,1]$. We call weak derivative the limit

$$\dot\nu_t=\lim_{\varepsilon\to 0}\frac{\nu_{t+\varepsilon}-\nu_t}{\varepsilon}, \tag{5}$$

provided this sequence converges in $M^1(\Sigma)$.

In the sequel, by distribution of $(l_t^n)_{t\in[0,1]}$ we will mean its distribution under the probability measure $\mathbb{P}^n(\cdot\,;\mu^n)$. It is an abuse of language, but there cannot be any confusion since the triangular array $((y_i^n)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ is fixed. Our first result is the following.

THEOREM 1. – *If $\mu^n \xrightarrow{w} \mu$ then $(l_t^n)_{t\in[0,1]}$ obeys a LDP on $D[[0,1],(M^+(\Sigma),\beta)]$ with good rate function*

$$I_\infty(\nu_\cdot,\mu)=\begin{cases}\int_0^1 H(\dot{\nu}_s \mid \mu)\,\mathrm{d}s & \text{if } \nu_\cdot\in\mathcal{AC}_\mu,\\ \infty & \text{elsewhere.}\end{cases} \tag{6}$$

Theorem 1 can be viewed as a LDP for the so-called microcanonical distributions. Simple microcanonical distributions are obtained from independent and identically distributed random variables $X_1,\ldots,X_n$ by conditioning on the value of a functional of their empirical measure. The question of interest is then whether or not there is convergence of the marginal distribution of X_1 under the conditional probability, when $n\to\infty$. For general background concerning microcanonical distributions we refer to Stroock and Zeitouni [18]. What we prove here is a LD result for the distribution of the contraction $(L_t^n=\frac{1}{n}\sum_{i=1}^{[nt]}\delta_{X_i})_{t\in[0,1]}$ of $X_1,\ldots,X_n$, when these random variables are n-exchangeable, under a strong conditioning.

Next, taking into account the fluctuations of the composition μ^n of the urn, we obtain in this case a more involved result. Let Q^n be the distribution of $L_1^n=\frac{1}{n}\sum_{i=1}^n\delta_{X_i^n}$ under $\mathbb{P}^n$. Note that this probability measure on $M^1(\Sigma)$ is also the distribution of $\frac{1}{n}\sum_{i=1}^n\delta_{Y_i^n}$ under P^n. Let $M^{1,n}(\Sigma)$ be the subset of $M^1(\Sigma)$ composed of all atomic measures $\frac{1}{n}\sum_{i=1}^n\delta_{x_i}$ for $(x_1,\ldots,x_n)\in\Sigma^n$ possibly with ties, and $\mathcal{AC}=\bigcup_{\mu\in M^1(\Sigma)}\mathcal{AC}_\mu$. Since

$$\mathbb{P}^n\big(L_\cdot^n\in A\big)=\int\limits_{M^{1,n}(\Sigma)}\mathbb{P}^n\big(l_\cdot^n\in A;\rho\big)\,Q^n(\mathrm{d}\rho) \tag{7}$$

for every borelian A of $D[[0,1],(M^+(\Sigma),\beta)]$, Theorem 1 tells us that $(L_t^n)_{t\in[0,1]}$ is a mixture of Large Deviation Systems (from now on abbreviated LDS), in the sense of Dawson and Gartner [3]. Hence, the announced LDP holds by virtue of a result due to Grunwald [13].

THEOREM 2. – *Suppose that L_1^n follows a LDP on $M^1(\Sigma)$ with good rate function J. Then $(L_t^n)_{t\in[0,1]}$ follows a LDP on $D[[0,1],(M^+(\Sigma),\beta)]$ with good rate function*

$$I(\nu_\cdot)=I_\infty(\nu_\cdot,\nu_1)+J(\nu_1)=\begin{cases}\int_0^1 H(\dot{\nu}_s\mid\nu_1)\,\mathrm{d}s+J(\nu_1) & \text{if } \nu_\cdot\in\mathcal{AC},\\ \infty & \text{elsewhere.}\end{cases} \tag{8}$$

Even in the simple case of binary valued finite exchangeable random variables there is no general result concerning the LD behavior of L_1^n. So Theorem 2 seems to be the best result that can be stated in this setting.

The paper is organized as follows. In Section 2 we consider a fixed triangular array $((y_i^n)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ of elements of Σ. Generalizing a technique from [5], we prove that if

$\mu^n = \frac{1}{n}\sum_{i=1}^n \delta_{y_i^n} \xrightarrow{w} \mu$ we have a LDP for $(l_{t_0}^n, \ldots, l_{t_d}^n)$ on $M^+(\Sigma)^{d+1}$, for all $d \in \mathbb{N}^*$ and all strictly ordered $(d+1)$-tuples $\underline{t} = (t_0 = 0 < t_1, \ldots, t_{d-1} < t_d = 1)$. We derive the LDP for $(l_t^n)_{t\in[0,1]}$ from the LDP for the finite-dimensional marginals $(l_{t_0}^n, \ldots, l_{t_d}^n)$ in Section 3. This result is obtained using a projective limit approach taken from [4]. In Section 4 we prove the identity (7) so that $(L_t^n)_{t\in[0,1]}$ is a mixture of LDS. Then we give the proof of Theorem 2, which is very close to the proof of Theorem 2.3 in [13]. Section 5 is devoted to applications of Theorem 2. We recover two classical examples of finite exchangeable random variables. We first consider the Curie–Weiss model, which is a well known toy model in statistical mechanics. Our analysis allows to consider both its microcanonical version (i.e., the uniform distribution on a set of allowed configurations), and its macrocanonical version (i.e., the classical Curie–Weiss model). These two aspects are connected via the principle of equivalence of ensembles. The Curie–Weiss model is a paradigm for both exchangeable random variables and LD problems as can be seen, for example, in the fact that its internal fluctuations are studied by means of a de Finetti representation by Papangelou in [16], and by the same author using LD techniques in [17]. Another classical example is given by infinite exchangeable sequences, where Theorem 2 allows us to extend easily the result of [9]. We also show that the LDP's for $(L_t^n)_{t\in[0,1]}$ where $X_1^n, \ldots, X_n^n$ are respectively given by sampling with and without replacement have closely related rate functions. This completes, in a way, a result of Baxter and Jain [2]. Our last example concerns the random permutation of a discrete time stochastic process. An n-tuple $(Y_1, \ldots, Y_n)$ is transformed into $(X_1^n, \ldots, X_n^n)$ by the mechanism presented above, i.e., $X_i^n = Y_{\sigma(i)}$ with $\sigma = \sigma^n$ a random permutation on $\{1, \ldots, n\}$, uniformly distributed and independent from $(Y_1, \ldots, Y_n)$. This appears to be a model for communication systems. A time-dependent signal Y^n is chopped into pieces of equal length $(Y_1, \ldots, Y_n)$ which are transmitted independently via different channels to the same destination. The signal is reconstructed according to the order of arrival into $X^n = (X_1^n, \ldots, X_n^n)$, whose LD behavior is given by Theorem 2.

2. Large deviations for finite marginals of $(l_t^n)_{t\in[0,1]}$

Let $((y_i^n)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ be a fixed triangular array of elements of Σ and let $d \in \mathbb{N}^*$ and $\underline{t} = (t_0 = 0 < t_1, \ldots, t_{d-1} < t_d = 1)$. Our objective in this section is to prove that if $\mu^n = \frac{1}{n}\sum_{i=1}^n \delta_{y_i^n} \xrightarrow{w} \mu$ then $(l_{t_0}^n, \ldots, l_{t_d}^n)$ follows a LDP on $M^+(\Sigma)^{d+1}$, with l_t^n as in (4). Fixing $(l_{t_0}^n, \ldots, l_{t_d}^n)$ is equivalent to choosing uniformly a partition of $(y_i^n)_{1\leqslant i\leqslant n}$ among those with d classes containing $[nt_j] - [nt_{j-1}]$ elements, for $1 \leqslant j \leqslant d$. In other words, we must associate to every y_i^n a value j, under the strong condition that $[nt_j] - [nt_{j-1}]$ items are associated to each j. First we relax the constraint on the cardinals of the d classes, and look for the LDP satisfied by the sequence of random measures

$$\mathcal{L}^n = \frac{1}{n}\sum_{i=1}^n \delta_{(y_i^n, N_i^n)}, \tag{9}$$

where the $((N_i^n)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ are independent random variables defined on a probability space $(\mathcal{Y}, \mathcal{F}, P)$, with values in a Polish space Γ, identically distributed according to a law λ. We will derive the LDP for $(l^n_{t_0}, \ldots, l^n_{t_d})$ from the latter result by conditioning on the values of N^n_i, thanks to a coupling.

LEMMA 1. – *The distribution of $\mathcal{L}^n$ under P obeys a LDP on $M^1(\Sigma \times \Gamma)$ endowed with the topology of weak convergence, with good rate function*

$$I_1(\nu, \mu, \lambda) = \begin{cases} H(\nu \mid \mu \otimes \lambda) & \text{if } \nu^{(1)} = \mu, \\ \infty & \text{otherwise,} \end{cases} \tag{10}$$

where $\nu^{(1)}$ stands for the first marginal of ν.

Proof. – Let $\phi \in C_b(\Sigma \times \Gamma)$, where we denote by $C_b(\Sigma \times \Gamma)$ the class of all real valued bounded continuous functions on $\Sigma \times \Gamma$. We have

$$\log E\left[\exp\left(n \int_{\Sigma\times\Gamma} \phi(u, v)\mathcal{L}^n(\mathrm{d}u \times \mathrm{d}v)\right)\right] = \log E\left[\exp \sum_{i=1}^{n} \phi\left(y^n_i, N^n_i\right)\right]$$
$$= \sum_{i=1}^{n} \log \int_{\Gamma} \exp\left(\phi\left(y^n_i, v\right)\right)\lambda(\mathrm{d}v),$$

then

$$\Lambda(\phi) = \lim_{n\to\infty} \frac{1}{n} \log E\left[\exp\left(n \int_{\Sigma\times\Gamma} \phi(u, v)\mathcal{L}^n(\mathrm{d}u \times \mathrm{d}v)\right)\right]$$
$$= \int_{\Sigma} \log\left(\int_{\Gamma} \exp\left(\phi(u, v)\right)\lambda(\mathrm{d}v)\right)\mu(\mathrm{d}u) < \infty.$$

Hence for all $k \in \mathbb{N}$ all $\phi_1, \ldots, \phi_k \in C_b(\Sigma \times \Gamma)$ and all $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ $\Lambda(\sum_{i=1}^k \lambda_i\phi_i)$ is finite and differentiable in $\lambda_1, \ldots, \lambda_k$ throughout $\mathbb{R}^k$. Whence, according to part *a*) of Corollary 4.6.11 in [6], $\mathcal{L}^n$ follows a LDP on $\mathcal{X}$, the algebraic dual of $C_b(\Sigma \times \Gamma)$, equipped with the $C_b(\Sigma \times \Gamma)$-topology, with good rate function

$$\Lambda^*(\nu) = \sup_{\phi\in C_b(\Sigma\times\Gamma)} \left\{\langle\phi, \nu\rangle - \Lambda(\phi)\right\},$$

where $\langle\cdot, \cdot\rangle$ stands as usual for

$$\langle\phi, \nu\rangle = \int_{\Sigma\times\Gamma} \phi \,\mathrm{d}\nu. \tag{11}$$

As $M^1(\Sigma \times \Gamma)$ is closed in $\mathcal{X}$ and $\Lambda^*(\nu) = \infty$ on $\mathcal{X}\backslash M^1(\Sigma \times \Gamma)$, $\mathcal{L}^n$ follows a LDP on $M^1(\Sigma \times \Gamma)$ equipped with the weak convergence topology, with good rate function Λ^*.

Let us identify Λ^*. From Theorem A.5.4 in [10] we know that every $\nu \in M^1(\Sigma \times \Gamma)$ can be written as $\nu(\mathrm{d}u \times \mathrm{d}v) = \nu^{(1)}(\mathrm{d}u) \otimes \rho(u, \mathrm{d}v)$, where ρ is a regular probability kernel.

First suppose that $\nu^{(1)} \neq \mu$. Then, there exists a $\phi \in C_b(\Sigma)$ such that $\int_\Sigma \phi(u)\nu^{(1)}(\mathrm{d}u) - \int_\Sigma \phi(u)\mu(\mathrm{d}u) = 1$, so for every $M > 0$ we define $\psi_M \in C_b(\Sigma \times \Gamma)$ by $\psi_M(u, v) = M\phi(u)$ such that

$$\int\limits_{\Sigma\times\Gamma} \psi_M(u,v)\nu(\mathrm{d}u\times\mathrm{d}v) - \int\limits_{\Sigma}\log\biggl(\int\limits_{\Gamma}\exp\bigl(\psi_M(u,v)\bigr)\lambda(\mathrm{d}v)\biggr)\mu(\mathrm{d}u)$$
$$= M\biggl(\int\limits_{\Sigma}\phi(u)\nu^{(1)}(\mathrm{d}u) - \int\limits_{\Sigma}\phi(u)\mu(\mathrm{d}u)\biggr) = M.$$

Whence we obtain in this case $\Lambda^*(\nu) = I_1(\nu,\mu,\lambda) = \infty$ by letting $M\to\infty$.

Now suppose that $\nu^{(1)} = \mu$. By virtue of Jensen's inequality, for any $\phi\in C_b(\Sigma\times\Gamma)$

$$\log\int\limits_{\Gamma}\int\limits_{\Sigma}\exp\bigl(\phi(u,v)\bigr)\lambda(\mathrm{d}v)\mu(\mathrm{d}u) \geqslant \int\limits_{\Sigma}\biggl(\log\int\limits_{\Gamma}\exp\bigl(\phi(u,v)\bigr)\lambda(\mathrm{d}v)\biggr)\mu(\mathrm{d}u).$$

Thus,

$$\int\limits_{\Sigma\times\Gamma}\phi(u,v)\nu(\mathrm{d}u\times\mathrm{d}v) - \log\int\limits_{\Gamma}\int\limits_{\Sigma}\exp\bigl(\phi(u,v)\bigr)\lambda(\mathrm{d}v)\mu(\mathrm{d}u)$$
$$\leqslant \int\limits_{\Sigma\times\Gamma}\phi(u,v)\nu(\mathrm{d}u\times\mathrm{d}v) - \int\limits_{\Sigma}\biggl(\log\int\limits_{\Gamma}\exp\bigl(\phi(u,v)\bigr)\lambda(\mathrm{d}v)\biggr)\mu(\mathrm{d}u).$$

Then, according to the definition of $H(\cdot\mid\cdot)$, we obtain $H(\nu\mid\mu\otimes\lambda)\leqslant\Lambda^*(\nu)$. So, if $H(\nu\mid\mu\otimes\lambda)=\infty$, we necessarily have $\Lambda^*(\nu) = I_1(\nu,\mu,\lambda)=\infty$. Otherwise, we can define

$$f(u,v) = \frac{\mathrm{d}(\mu\otimes\rho)}{\mathrm{d}(\mu\otimes\lambda)} = \frac{\mathrm{d}\rho}{\mathrm{d}\lambda}\,\mu\otimes\lambda \quad \text{a.e.}$$

For every $\phi\in C_b(\Sigma\times\Gamma)$

$$H\bigl(\rho(u,\cdot)\mid\lambda\bigr) \geqslant \int\limits_{\Gamma}\phi(u,v)\rho(u,\mathrm{d}v) - \log\int\limits_{\Gamma}\exp\bigl(\phi(u,v)\bigr)\lambda(\mathrm{d}v) \quad \mu \text{ a.e.},$$

hence

$$\int\limits_{\Sigma}H\bigl(\rho(u,\cdot)|\lambda\bigr)\mu(\mathrm{d}u) \geqslant \int\limits_{\Sigma\times\Gamma}\phi(u,v)\nu(\mathrm{d}u\times\mathrm{d}v) - \int\limits_{\Sigma}\log\biggl(\int\limits_{\Gamma}\exp\bigl(\phi(u,v)\bigr)\lambda(\mathrm{d}v)\biggr)\mu(\mathrm{d}u),$$

so $\int_\Sigma H(\rho(u,\cdot)\mid\lambda)\mu(\mathrm{d}u)\geqslant\Lambda^*(\nu)$.

But, according to Fubini's theorem

$$\int\limits_{\Sigma}H\bigl(\rho(u,\cdot)|\lambda\bigr)\mu(\mathrm{d}u) = \int\limits_{\Sigma}\biggl(\int\limits_{\Gamma}\frac{\mathrm{d}\rho}{\mathrm{d}\lambda}\log\frac{\mathrm{d}\rho}{\mathrm{d}\lambda}\,\mathrm{d}\lambda\biggr)\mathrm{d}\mu$$
$$= \int\limits_{\Sigma\times\Gamma}\frac{\mathrm{d}(\mu\otimes\rho)}{\mathrm{d}(\mu\otimes\lambda)}\log\frac{\mathrm{d}(\mu\otimes\rho)}{\mathrm{d}(\mu\otimes\lambda)}\,\mathrm{d}(\mu\otimes\lambda)$$
$$= H(\nu\mid\mu\otimes\lambda).$$

so $H(\nu\mid\mu\otimes\lambda)\geqslant\Lambda^*(\nu)$ and then $\lambda^*(\nu) = I_1(\nu,\mu,\lambda)$. $\square$

We proceed now to the identification of each N_i^n $(1 \leqslant i \leqslant n)$ with an element of a random partition in d classes of $(y_i^n)_{1\leqslant i\leqslant n}$. We suppose that $\Gamma = \{1, \ldots, d\}$, that the N_i^n are distributed according to $\lambda(j) = t_j - t_{j-1} =: \Delta_j t$, and we define the continuous and injective map

$$\begin{aligned} F : M^1(\Sigma \times \Gamma) &\longrightarrow M^+(\Sigma)^d \\ \nu(\cdot,\cdot) &\longmapsto \big(\nu(\cdot,\{1\}), \nu(\cdot,\{1,2\}), \ldots, \nu(\cdot,\Gamma)\big). \end{aligned} \tag{12}$$

For every $n \in \mathbb{N}$ we set

$$\mathcal{S}^n = F \circ \mathcal{L}^n, \tag{13}$$

with $\mathcal{L}^n$ as in (9). The vector of random measures $\mathcal{S}^n$ is defined on $(\mathcal{Y}, \mathcal{F}, P)$ as in Lemma 1. An element $\nu = (\nu_i)_{i\in\Gamma}$ of $M^+(\Sigma)^d$ is said to be *increasing* when $\nu_i(A) \geqslant \nu_j(A)$ for all $A \in \mathcal{B}_\Sigma$ and all $i, j \in \Gamma$ such that $i \geqslant j$. For these elements of $M^+(\Sigma)^d$ we denote by $\Delta_i\nu$ the positive measure $\nu_i - \nu_{i-1}$, with $\nu_0 = 0$.

COROLLARY 1. – *The distribution of $\mathcal{S}^n$ under P obeys a LDP on $M^+(\Sigma)^d$ equipped with the product topology of weak convergence, with good rate function*

$$I_2(\nu,\mu,\underline{t}) = \begin{cases} \displaystyle\sum_{i=1}^d \Delta_i\nu(\Sigma) H\left(\frac{\Delta_i\nu}{\Delta_i\nu(\Sigma)}\Big|\mu\right) & \\ \displaystyle\qquad + \sum_{i=1}^d \Delta_i\nu(\Sigma)\log\frac{\Delta_i\nu(\Sigma)}{\Delta_i t} & \textit{if} \begin{cases} \nu \textit{ is increasing,} \\ \nu_d = \mu, \end{cases} \\ \infty & \textit{elsewhere.} \end{cases} \tag{14}$$

Proof. – Let $\mathcal{M} = F(M^1(\Sigma \times \Gamma)) = \{\nu \in M^+(\Sigma)^d,\ \nu \text{ increasing and } \nu_d(\Sigma) = 1\}$. Since F is continuous and injective, we deduce from Lemma 1 that $\mathcal{S}^n$ follows a LDP on $M^+(\Sigma)^d$ endowed with the product topology of weak convergence, with good rate function

$$\bar{I}_2(\nu,\mu,\underline{t}) = \begin{cases} I_1(\nu^*,\mu,\lambda) & \text{if } \nu \in \mathcal{M} \text{ and } \nu = F(\nu^*), \\ \infty & \text{elsewhere,} \end{cases}$$

where I_1 is the rate function defined in (10).

If $\nu \notin \mathcal{M}$, $\bar{I}_2(\nu,\mu,\underline{t}) = I_2(\nu,\mu,\underline{t}) = \infty$. Let $\nu \in \mathcal{M}$. Then we have $\nu_d = \nu^{*(1)}$, the first marginal of ν^* and if $\nu_d \neq \mu$, $I_2(\nu,\mu,\underline{t}) = \bar{I}_2(\nu,\mu,\underline{t}) = \infty$. If $\nu_d = \mu$ then ν^* is absolutely continuous w.r.t. $\mu \otimes \lambda$ and

$$\begin{aligned} \bar{I}_2(\nu,\mu,\underline{t}) &= I_1(\nu^*,\mu,\lambda) \\ &= H(\nu^* \mid \mu\otimes\lambda) \\ &= \sum_{i=1}^d \int_\Sigma \nu^*(\mathrm{d}y,i)\log\frac{\nu^*(\mathrm{d}y,i)}{\mu(\mathrm{d}y)\Delta_i t} \\ &= \sum_{i=1}^d \Delta_i\nu(\Sigma)\int_\Sigma \frac{\Delta_i\nu(\mathrm{d}y)}{\Delta_i\nu(\Sigma)}\log\frac{\Delta_i\nu(\mathrm{d}y)/\Delta_i\nu(\Sigma)}{\mu(\mathrm{d}y)\Delta_i t/\Delta_i\nu(\Sigma)} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{d} \Delta_i \nu(\Sigma) H\left(\frac{\Delta_i \nu}{\Delta_i \nu(\Sigma)}\Big|\mu\right) + \sum_{i=1}^{d} \Delta_i \nu(\Sigma) \log \frac{\Delta_i \nu(\Sigma)}{\Delta_i t}\\
&= I_2(\nu, \mu, \underline{t}).
\end{aligned}
$$

Hence we obtain the rate function of the LDP satisfied by $\mathcal{S}^n$. □

Next we define a coupling procedure that allows us to derive from $\mathcal{S}^n$ a random variable with the same law as $(l_{t_1}^n, \dots, l_{t_d}^n)$. Let U_j^n be the number of j-valued N_i^n ($j \in \{1, \dots, d\}$), and T_n be the typical event $T_n = \bigcap_{j=1}^d \{U_j^n = [nt_j] - [nt_{j-1}]\}$. For every $n \in \mathbb{N}$ we define $(\tilde{N}_i^n)_{1 \leqslant i \leqslant n}$ from $(N_i^n)_{1 \leqslant i \leqslant n}$ in the following way:

- If U_1^n is greater than $[nt_1]$, we choose randomly $U_1^n - [nt_1]$ i's among the ones with $N_i^n = 1$, and we change the value 1 on these i's to the value 2.
- If U_1^n is less than $[nt_1]$, we choose uniformly $[nt_1] - U_1^n$ indices among those such that $N_i^n = 2$, and we change the associated N_i^n into 1. If there are not enough i's such that $N_i^n = 2$, we choose the needed indices among those with $N_i^n = 3$.

We call $\bar{N}_{i,1}^n \in \{1, \dots, d\}$ the random variables resulting from this first step of the procedure. Now we define the random variables $\bar{N}_{i,2}^n \in \{1, \dots, d\}$ resulting from the second step in the same way:

- If the number of i's with $\bar{N}_{i,1}^n = 2$ is greater than $[nt_2] - [nt_1]$, we choose uniformly the indices in excess, and we change the value 2 on these i's to the value 3.
- If the number of $\bar{N}_{i,1}^n = 2$ is less than $[nt_2] - [nt_1]$, we complete it by choosing uniformly indices among those such that $\bar{N}_{i,1}^n = 3$. If there are not enough i's such that $N_i^n = 3$, we choose the needed indices among those such that $N_i^n = 4$.

We carry on up to $d - 1$, and we set $\bar{N}_{i,j}^n \in \{1, \dots, d\}$ for the ith random variable at the jth step of the coupling procedure. We put $(\bar{N}_{i,0}^n)_{1 \leqslant i \leqslant n} = (N_i^n)_{1 \leqslant i \leqslant n}$ and we define the $(\tilde{N}_i^n)_{1 \leqslant i \leqslant n}$ by $\tilde{N}_i^n = \bar{N}_{i,d-1}^n$. For every $n \in \mathbb{N}$ we note

$$
\tilde{\mathcal{L}}^n = \frac{1}{n} \sum_{i=1}^{n} \delta_{(y_i^n, \tilde{N}_i^n)}, \tag{15}
$$

and

$$
\widetilde{\mathcal{S}}^n = F \circ \tilde{\mathcal{L}}^n, \tag{16}
$$

with F as in (12).

LEMMA 2. – *For every $n \in \mathbb{N}$ the law of $\widetilde{\mathcal{S}}^n$ is the law of $\mathcal{S}^n$ conditioned on T_n, and for every measurable $B \subset M^+(\Sigma)^d$ we have*

$$
P\big(\widetilde{\mathcal{S}}^n \in B\big) = \mathbb{P}^n\big(\big(l_{t_1}^n, \dots, l_{t_d}^n\big) \in B;\ \mu^n\big).
$$

Proof. – Even if the random variables $\widetilde{\mathcal{S}}^n$ and $(l_{t_1}^n, \dots, l_{t_d}^n)$ are defined on different probability spaces, their distribution on $M^+(\Sigma)^d$ have the same finite support A^n, and it is also the support of the distribution of $\mathcal{S}^n$ conditioned on T_n. Since $(x_1^n, \dots, x_n^n)$ results from a sampling without replacement all possible $(l_{t_1}^n, \dots, l_{t_d}^n)$ are equally-likely, thus for every $\rho \in A^n$

$$\mathbb{P}^n\big((l^n_{t_1}, \dots, l^n_{t_d}) = \rho; \mu^n\big) = \frac{1}{|A^n|}.$$

The cardinal of A^n might not be

$$\prod_{i=1}^{d} \binom{n - [nt_{i-1}]}{[nt_i] - [nt_{i-1}]}$$

because of possible ties among $(y^n_1, \dots, y^n_n)$. In the same time, as the law of $(N^n_1, \dots, N^n_n)$ conditioned on T_n is uniform on its support, for every $\rho \in A^n$

$$P\big(\mathcal{S}^n = \rho \mid T_n\big) = \frac{1}{|A^n|}.$$

Hence it is then sufficient to prove that $\widetilde{\mathcal{S}}^n$ is uniformly distributed on A^n. For all $\rho, \gamma \in \mathrm{Im}(\widetilde{\mathcal{S}}^n)$ there are $\underline{u} = (u_i)_{1\leqslant i\leqslant n}$ and $\underline{v} = (v_i)_{1\leqslant i\leqslant n}$ such that we have $\{\widetilde{\mathcal{S}}^n = \rho\} = \{\tilde{N}^n_1 = u_1, \dots, \tilde{N}^n_n = u_n\}$ and $\{\widetilde{\mathcal{S}}^n = \gamma\} = \{\tilde{N}^n_1 = v_1, \dots, \tilde{N}^n_d = v_n\}$, and there is a permutation σ on $\{1, \dots, n\}$ such that for all i $u_i = v_{\sigma(i)}$. Hence, $P(\widetilde{\mathcal{S}}^n = \rho) = P(\widetilde{\mathcal{S}}^n = \gamma)$ if and only if $(\tilde{N}^n_i)_{1\leqslant i\leqslant n}$ is n-exchangeable. In order to prove it we introduce the following notations:

- $V^{\underline{v}}_{\underline{u}}(j)$ stands for the event:
  "The jth step of the coupling procedure changes $(\bar{N}^n_{i,j-1})_{1\leqslant i\leqslant n} = \underline{u}$ to $(\bar{N}^n_{i,j})_{1\leqslant i\leqslant n} = \underline{v}$".
- For all $1 \leqslant i \leqslant n$ and for all $1 \leqslant q \leqslant d$ we call $Y^q_i = (\bar{N}^n_{i,0}, \dots, \bar{N}^n_{i,q-1}) \in \{1, \dots, d\}^q$ the random vector that records the values associated to i during the procedure.

Note that what matters in $V^{\underline{v}}_{\underline{u}}(j)$ is the number of k-valued u_i's and v_i's in $\underline{u}$ and $\underline{v}$ for each $k \in \{j, \dots, d\}$, not the value of each u_i and v_i. Hence, for every permutation σ on $\{1, \dots, n\}$ we have $P(V^{\underline{v}}_{\underline{u}}(j)) = P(V^{\sigma(\underline{v})}_{\sigma(\underline{u})}(j))$, where $\sigma(\underline{u}) = (u_{\sigma(1)}, \dots, u_{\sigma(n)})$.

We prove by induction on q that for every $1 \leqslant q \leqslant d$, $(Y^q_i)_{1\leqslant i\leqslant n}$ is n-exchangeable. For $q = 1$, the $(\bar{N}^n_{i,0})_{1\leqslant i\leqslant n}$ are independent and identically distributed, whence $(Y^1_i)_{1\leqslant i\leqslant n}$ is n-exchangeable. Suppose the property holds for a fixed q $(1 \leqslant q \leqslant d-1)$: $(Y^q_i = (\bar{N}^n_{i,0}, \dots, \bar{N}^n_{i,q-1}))_{1\leqslant i\leqslant n}$ is n-exchangeable. Let $(u^j_i) \in \mathcal{M}_{n,q+1}(\Gamma)$, we denote by $\underline{u}_i$ its ith row and by $\underline{u}^j$ its jth column. For every permutation σ on $\{1, \dots, n\}$

$$\begin{aligned}
&P\big(Y^{q+1}_i = \underline{u}_i,\ 1 \leqslant i \leqslant n\big)\\
&\quad = P\big(\bar{N}^n_{i,0} = u^0_i, \dots, \bar{N}^n_{i,q} = u^q_i, 1 \leqslant i \leqslant n\big)\\
&\quad = P\big(\bar{N}^n_{i,0} = u^0_i, \dots, \bar{N}^n_{i,q-1} = u^{q-1}_i, V^{\underline{u}^q}_{\underline{u}^{q-1}}(q),\ 1 \leqslant i \leqslant n\big)\\
&\quad = P\big(V^{\underline{u}^q}_{\underline{u}^{q-1}}(q) | \bar{N}^n_{i,q-1} = u^{q-1}_i,\ 1 \leqslant i \leqslant n\big) P\big(\bar{N}^n_{i,0} = u^0_i, \dots, \bar{N}^n_{i,q-1} = u^{q-1}_i,\ 1 \leqslant i \leqslant n\big)\\
&\quad = P\big(V^{\sigma(\underline{u}^{q-1})}_{\sigma(\underline{u}^{q-1})}(q)\big) P\big(\bar{N}^n_{\sigma(i),0} = u^0_i, \dots, \bar{N}^n_{\sigma(i),q-1} = u^{q-1}_i,\ 1 \leqslant i \leqslant n\big)\\
&\quad = P\big(Y^{q+1}_{\sigma(i)} = \underline{u}_i, 1 \leqslant i \leqslant n\big).
\end{aligned}$$

Hence we obtain that $(Y_i^{q+1})_{1\leqslant i\leqslant n}$ is n-exchangeable, so $(Y_i^d)_{1\leqslant i\leqslant n}$ is also n-exchangeable, and in particular $(\tilde{N}_i^n)_{1\leqslant i\leqslant n}$ is. □

The last two results lead to the announced crucial lemma.

LEMMA 3. – *$(l_{t_0}^n,\dots,l_{t_d}^n)$ obeys a LDP on $M^+(\Sigma)^{d+1}$ endowed with the product topology of weak convergence, with good rate function*

$$I_3(\nu,\mu,\underline{t})=\begin{cases}\displaystyle\sum_{i=1}^d \Delta_i t H\left(\frac{\Delta_i\nu}{\Delta_i t}\Big|\mu\right) & \textit{if}\begin{cases}\nu \textit{ is increasing},\\ \nu_d=\mu,\\ \forall i\in\{0,\dots,d\}\ \nu_i(\Sigma)=t_i,\end{cases}\\ \infty & \textit{elsewhere}.\end{cases}\tag{17}$$

Proof. – Since for every $n\in\mathbb{N}$ $(l_{t_0}^n,\dots,l_{t_d}^n)\in\{0\}\times M^+(\Sigma)^d$ which is a closed subset of $M^+(\Sigma)^{d+1}$, it is sufficient to prove that $(l_{t_1}^n,\dots,l_{t_d}^n)$ follows a LDP on $M^+(\Sigma)^d$ with good rate function

$$\bar{I}_3(\nu,\mu,\underline{t})=\begin{cases}\displaystyle\sum_{i=1}^d \Delta_i t H\left(\frac{\Delta_i\nu}{\Delta_i t}\Big|\mu\right) & \text{if}\begin{cases}\nu \text{ is increasing},\\ \nu_d=\mu,\\ \forall i\in\{1,\dots,d\}\ \nu_i(\Sigma)=t_i,\end{cases}\\ \infty & \text{elsewhere}.\end{cases}$$

We first prove the upper bound of this LDP. Let A be a closed part of $M^+(\Sigma)^d$. For all $\varepsilon>0$ we note $R_\varepsilon=\{\nu\in M^+(\Sigma)^d,\ \sup_i|\nu_i(\Sigma)-t_i|\leqslant\varepsilon\}$. For ε fixed and large enough n $\{\mathcal{S}^n\in A\}\cap T_n\subset\{\mathcal{S}^n\in A\cap R_\varepsilon\}$. Then, according to Corollary 1

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\log P\bigl(\{\mathcal{S}^n\in A\}\cap T_n\bigr)&\leqslant\limsup_{n\to\infty}\frac{1}{n}\log P\bigl(\mathcal{S}^n\in A\cap R_\varepsilon\bigr)\\&\leqslant-\inf_{A\cap R_\varepsilon}I_2(\nu,\mu,\underline{t}),\end{aligned}$$

I_2 as in (14). Since I_2 is a good rate function

$$\lim_{\varepsilon\to0}\inf_{A\cap R_\varepsilon}I_2(\nu,\mu,\underline{t})=\inf_{A\cap R_0}I_2(\nu,\mu,\underline{t})=\inf_A\bar{I}_3(\nu,\mu,\underline{t}).$$

Furthermore

$$P(T_n)=\prod_{i=1}^d\binom{n-[nt_{i-1}]}{[nt_i]-[nt_{i-1}]}(\Delta_i t)^{[nt_i]-[nt_{i-1}]},$$

so we obtain $\liminf_{n\to\infty}\frac{1}{n}\log P(T_n)=0$. Thus, according to Lemma 2

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}^n\bigl((l_{t_1}^n,\dots,l_{t_d}^n)\in A;\mu^n\bigr)\\&\quad=\limsup_{n\to\infty}\frac{1}{n}\log P\bigl(\widetilde{\mathcal{S}}^n\in A\bigr)\\&\quad\leqslant\limsup_{n\to\infty}\frac{1}{n}\log P\bigl(\{\mathcal{S}^n\in A\}\cap T_n\bigr)-\liminf_{n\to\infty}\frac{1}{n}\log P(T_n)\\&\quad\leqslant-\inf_{A\cap R_\varepsilon}I_2(\nu,\mu,\underline{t}).\end{aligned}$$

Hence we have the upper bound of the LDP for $(l^n_{t_1}, \dots, l^n_{t_d})$ by letting $\varepsilon \to 0$.

Next we prove the lower bound of the LDP. Let us recall that the dual-bounded-Lipschitz metric β defined on $M^+(\Sigma)$ by

$$\beta(\rho, \nu) = \sup\left\{\left|\int_\Sigma f\,\mathrm{d}\rho - \int_\Sigma f\,\mathrm{d}\nu\right|,\ f \in C_b(\Sigma),\ \|f\|_\infty + \|f\|_L \leqslant 1\right\} \tag{18}$$

with

$$\|f\|_\infty = \sup_{x\in\Sigma}\left|f(x)\right| \quad \text{and} \quad \|f\|_L = \sup_{x,y\in\Sigma, x\neq y}\left|\frac{f(x)-f(y)}{d(x,y)}\right|$$

coincides with the weak convergence topology (see Appendix A.1 in [4]). We denote by β_d the supremum metric on $M^+(\Sigma)^d$ associated to β. Let C be an open subset of $M^+(\Sigma)^d$, and $\nu \in C$ be such that $\bar{I}_3(\nu, \mu, \underline{t}) < \infty$. Since for all $i \in \{1, \dots, d\}$, $\nu_i(\Sigma) = t_i$, there exists, for all $n \in \mathbb{N}$, a $\nu^n \in M^+(\Sigma)^d$ with $\nu^n_i(\Sigma) = \frac{[nt_i]}{n}$ such that the sequence $(\nu^n)_{n\in\mathbb{N}}$ satisfies $\nu^n \xrightarrow{w} \nu$. For every $j \in \{1, \dots, d\}$ we define

$$D_j = \left\{i \in \{1, \dots, n\},\ \left(N^n_i \leqslant j \text{ and } \tilde{N}^n_i > j\right) \text{ or } \left(N^n_i > j \text{ and } \tilde{N}^n_i \leqslant j\right)\right\},$$

and for all f with $\|f\|_\infty \leqslant 1$ we have

$$\begin{aligned}
\left|\int_\Sigma f\,\mathrm{d}\mathcal{S}^n_j - \int_\Sigma f\,\mathrm{d}\widetilde{\mathcal{S}}^n_j\right| &= \frac{1}{n}\left|\sum_{y^n_i:\tilde{N}^n_i\leqslant j}\left(y^n_i\right) - \sum_{y^n_i\,:\,N^n_i\leqslant j} f\left(y^n_i\right)\right| \\
&\leqslant \frac{1}{n}\sum_{i\in D_j}\left|f\left(y^n_i\right)\right| \leqslant \frac{|D_j|}{n} \\
&= \left|\mathcal{S}^n_j(\Sigma) - \frac{[nt_j]}{n}\right| \leqslant \beta_d\left(\mathcal{S}^n, \nu^n\right).
\end{aligned}$$

Hence $\beta_d(\mathcal{S}^n, \widetilde{\mathcal{S}}^n) \leqslant \beta_d(\mathcal{S}^n, \nu^n)$. Combining the preceding inequality and the triangular inequality we obtain, for all $\delta > 0$ and n large enough

$$\begin{aligned}
P\left(\beta_d\left(\widetilde{\mathcal{S}}^n, \nu\right) < 5\delta\right) &\geqslant P\left(\beta_d\left(\mathcal{S}^n, \widetilde{\mathcal{S}}^n\right) < 2\delta, \beta_d\left(\mathcal{S}^n, \nu^n\right) < 2\delta\right) \\
&= P\left(\beta_d\left(\mathcal{S}^n, \nu^n\right) < 2\delta\right) \\
&\geqslant P\left(\beta_d\left(\mathcal{S}^n, \nu\right) < \delta\right).
\end{aligned}$$

Let $\delta > 0$ be such that $B_{\beta_d}(\nu, 5\delta) \subset C$, where B_{β_d} stands for an open ball defined with the metric β_d. Corollary 1 and Lemma 2 tell us that

$$\begin{aligned}
\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}^n\left(\left(l^n_{t_1}, \dots, l^n_{t_d}\right) \in C; \mu^n\right) &= \liminf_{n\to\infty}\frac{1}{n}\log P\left(\widetilde{\mathcal{S}}^n \in C\right) \\
&\geqslant \liminf_{n\to\infty}\frac{1}{n}\log P\left(\widetilde{\mathcal{S}}^n \in B_{\beta_d}(\nu, 5\delta)\right) \\
&\geqslant \liminf_{n\to\infty}\frac{1}{n}\log P\left(\mathcal{S}^n \in B_{\beta_d}(\nu, \delta)\right) \\
&\geqslant -I_2(\nu, \mu, \underline{t}) = -\bar{I}_3(\nu, \mu, \underline{t}).
\end{aligned}$$

Hence we get the lower bound of the LDP followed by $(l^n_{t_1}, \dots, l^n_{t_d})$.

Last we prove that $\bar{I}_3$ is a good rate function. For every $0 \leqslant \alpha < \infty$

$$\begin{aligned}\phi^{\bar{I}_3}_\alpha &= \{\nu \in M^+(\Sigma)^d,\ \bar{I}_3(\nu, \mu, \underline{t}) \leqslant \alpha\} \\ &= \{\nu \in M^+(\Sigma)^d,\ I_2(\nu, \mu, \underline{t}) \leqslant \alpha\} \cap \{\nu \in M^+(\Sigma),\ \Delta_i \nu(\Sigma) = \Delta_i t\}.\end{aligned}$$

Since I_2 is a good rate function, $\phi^{\bar{I}_3}_\alpha$ is the intersection of a compact and a closed subset in the weak convergence topology. Hence it is compact and $\bar{I}_3$ is a good rate function. □

3. Large deviations for the process $(l^n_t)_{t\in[0,1]}$

Our aim in this section is to derive the LDP for $(l^n_t)_{t\in[0,1]}$ from the LDP for the finite-dimensional marginals $(l^n_{t_0}, \dots, l^n_{t_d})$. We use a projective limit approach, as in the proof of Theorem 1 in [4]. Since our setting, and then our proof, is slightly different, we give it completely for the sake of clarity.

Let $C[[0,1], (M^+(\Sigma), \beta)]$ be the space of all maps that are continuous from $[0,1]$ to $M^+(\Sigma)$. Unless explicitly stated otherwise, it is equipped with the uniform metric β_∞ as in (2). We still consider a fixed triangular array $((y^n_i)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ of elements of Σ which composition given by $(\mu^n = \frac{1}{n}\sum_{i=1}^n \delta_{y^n_i})_{n\in\mathbb{N}}$ satisfies $\mu^n \xrightarrow{w} \mu$. We define by

$$\bar{l}^n_t = l^n_t + \left(t - \frac{[nt]}{n}\right)\delta_{x^n_{[nt]+1}} \tag{19}$$

the linear interpolation $(\bar{l}^n_t)_{t\in[0,1]}$ of $(l^n_t)_{t\in[0,1]}$, l^n_t being as in (4). Remark that $(\bar{l}^n_t)_{t\in[0,1]} \in C[[0,1], (M^+(\Sigma), \beta)]$. Let us recall that we consider the distribution of $(\bar{l}^n_t)_{t\in[0,1]}$ and $(l^n_t)_{t\in[0,1]}$ under the probability measure $\mathbb{P}^n(\cdot\,; \mu^n)$ associated to the sampling without replacement among $(y^n_1, \dots, y^n_n)$. First we prove a LDP for $(\bar{l}^n_t)_{t\in[0,1]}$ for which we give an explicit rate function. We need to consider the linear interpolation because it is the only way to pass from results on the pointwise convergence topology to results on the uniform convergence topology. Since $(l^n_t)_{t\in[0,1]}$ and $(\bar{l}^n_t)_{t\in[0,1]}$ are exponentially equivalent, we deduce the LDP satisfied by $(l^n_t)_{t\in[0,1]}$ from the preceding result.

LEMMA 4. – 1. *$(l^n_t)_{t\in[0,1]}$ and $(\bar{l}^n_t)_{t\in[0,1]}$ are exponentially equivalent on $D[[0,1], (M^+(\Sigma), \beta)]$.*

2. *$(\bar{l}^n_t)_{t\in[0,1]}$ is exponentially tight on the Polish space $(C[[0,1], (M^+(\Sigma), \beta)], \beta_\infty)$.*

Proof. – 1. With probability 1, we have

$$\begin{aligned}\beta_\infty(l^n_\cdot, \bar{l}^n_\cdot) &= \sup_{t\in[0,1]} \beta\left(l^n_t, l^n_t + \left(t - \frac{[nt]}{n}\right)\delta_{x^n_{[nt]+1}}\right) \\ &\leqslant \sup_{t\in[0,1]}\left(t - \frac{[nt]}{n}\right) \leqslant \frac{1}{n}.\end{aligned}$$

Then $(l^n_t)_{t\in[0,1]}$ and $(\bar{l}^n_t)_{t\in[0,1]}$ are exponentially equivalent on $D[[0,1], (M^+(\Sigma), \beta)]$.

2. According to Lemma 3, for every fixed $t \in [0,1]$, $\bar{l}^n_t$ follows a LDP on the Polish space $(M^+(\Sigma), \beta)$ with a good rate function. Hence it is exponentially tight.

Furthermore

$$\beta\left(\bar{l}_t^n, \bar{l}_s^n\right) \leqslant \frac{[nt]-[ns]}{n},$$

so we can conclude thanks to Appendix A.2 in [4]. $\square$

Let G be the set of all the subdivisions $0 = t_0 < \cdots < t_d = 1$ of $[0,1]$. We define on G the partial order $i = (s_0, \ldots, s_p) \leqslant j = (t_0, \ldots, t_q)$ if and only if for all $s_u \in i$ there is a $t_v \in j$ such that $s_u = t_v$, which makes G a right-filtering set. For $i = (s_0, \ldots, s_p) \leqslant j = (t_0, \ldots, t_q)$ we define $p_{ij}(\nu_{t_0}, \ldots, \nu_{t_q}) = (\nu_{s_0}, \ldots, \nu_{s_p})$. Endowing $M^+(\Sigma)^{|j|}$ with the product topology associated to β makes $(M^+(\Sigma)^{|j|}, p_{ij})_{i \leqslant j}$ a projective system which projective limit is $\mathcal{E} = \{\nu : [0,1] \to M^+(\Sigma)\}$ equipped with the topology of pointwise convergence. For every $j = (t_0, \ldots, t_q) \in G$, we note p_j the canonical projection of $\mathcal{E}$ on $M^+(\Sigma)^{|j|}$, and we define on $\mathcal{E}$ the map $I_j(\nu_\cdot, \mu) = I_3(p_j \nu_\cdot, \mu, j)$, with I_3 as in (17). Next we prove a LDP for $(\bar{l}_t^n)_{t \in [0,1]}$ in $\mathcal{E}$.

LEMMA 5. – $(\bar{l}_t^n)_{t \in [0,1]}$ *follows a LDP on* $\mathcal{E}$ *with good rate function*

$$I_\infty(\nu_\cdot, \mu) = \sup_{j \in G} I_j(\nu_\cdot, \mu). \tag{20}$$

Proof. – Since $(\bar{l}_t^n)_{t \in [0,1]}$ and $(l_t^n)_{t \in [0,1]}$ are exponentially equivalent on $D[[0,1], (M^+(\Sigma), \beta)]$, we deduce from Lemma 3 that for every $j \in G p_j(\bar{l}_\cdot^n)$ follows a LDP on $M^+(\Sigma)^{|j|}$ with good rate function $I_j(\nu_\cdot, \mu)$. Hence, according to Dawson–Gartner's theorem, $(\bar{l}_t^n)_{t \in [0,1]}$ obeys a LDP on $\mathcal{E}$ with good rate function $I_\infty(\nu_\cdot, \mu) = \sup_{j \in G} I_j(\nu_\cdot, \mu)$. $\square$

We recall that $\mathcal{AC}_\mu$ is the space of all maps $\nu_t : [0,1] \to M^+(\Sigma)$ such that:

1. $\nu_t - \nu_s \in M^+(\Sigma)$ is of total mass $t - s$ for all $0 \leqslant s \leqslant t$.
2. $\nu_0 = 0$ and $\nu_1 = \mu$.
3. $\nu_\cdot$ possesses a weak derivative for almost every $t \in [0,1]$ as defined in (5).

The following result gives an explicit expression of $I_\infty(\cdot, \mu)$ on $D[[0,1], (M^+(\Sigma), \beta)]$.

LEMMA 6. – 1. *For every* $\nu_\cdot \in D[[0,1], (M^+(\Sigma), \beta)]$, *if* $I_\infty(\nu_\cdot, \mu) < \infty$ *then* $\nu_\cdot \in \mathcal{AC}_\mu$.

2. *For all* $\nu_\cdot \in \mathcal{AC}_\mu$, $I_\infty(\nu_\cdot, \mu) = \int_0^1 H(\dot{\nu}_s \mid \mu)\, \mathrm{d}s$.

Proof. – 1. Let $\nu_\cdot \in D[[0,1], (M^+(\Sigma), \beta)]$ be such that $I_\infty(\nu_\cdot, \mu) < \infty$. For every $j = (0, s, t, 1)$ we necessarily have $I_j(\nu_\cdot, \mu) < \infty$. Hence $\nu_t - \nu_s \in M^+(\Sigma)$ is of total mass $t - s$, $\nu_0 = 0$ and $\nu_1 = \mu$.

As $I_\infty(\nu_\cdot, \mu) < \infty$, we have for all $j = (t_0, \ldots, t_d) \in G$

$$I_j(\nu_\cdot, \mu) = \sum_{i=1}^d (t_i - t_{i-1}) H\left(\frac{\nu_{t_i} - \nu_{t_{i-1}}}{t_i - t_{i-1}} \middle| \mu\right).$$

For every $n \in \mathbb{N}$ we define the process $g_n : [0,1] \to M^1(\Sigma)$ by

$$g_n(t) = 2^n \left[\nu_{\frac{[2^n t]+1}{2^n}} - \nu_{\frac{[2^n t]}{2^n}}\right].$$

We get

$$I_\infty(\nu_\cdot,\mu) \geqslant \sum_{i=1}^{2^n} 2^{-n} H\big(2^n\big(\nu_{\frac{i}{2^n}} - \nu_{\frac{i-1}{2^n}}\big) \mid \mu\big) = \int_0^1 H\big(g_n(t) \mid \mu\big)\mathrm{d}t$$

and, $H(\cdot \mid \mu)$ being convex,

$$\int_{\frac{i-1}{2^n}}^{\frac{i}{2^n}} H\big(g_{n+1}(t) \mid \mu\big)\,\mathrm{d}t \geqslant H\left(g_n\left(\frac{i}{2^n}\right)\Big|\mu\right)$$

for all $i = 1,\dots,2^n$. The previous inequality tells us that the sequence of real valued random variables $(H(g_n \mid \mu))_{n\in\mathbb{N}}$ defined on $[0,1]$ endowed with the Lebesgue measure and the dyadic filtration $\mathcal{F}_n = \sigma([\frac{j-1}{2^n},\frac{j}{2^n}),\ 1 \leqslant j \leqslant 2^n)$ is a submartingale. Since we have $\sup_{n\in\mathbb{N}} \int_0^1 H(g_n(t)\mu)\,\mathrm{d}t < \infty$, we know that

$$b(t) = 1 + \limsup_{n\to\infty} H\big(g_n(t) \mid \mu\big) < \infty$$

for a.e. $t \in [0,1]$ by virtue of Doob's theorem. But, for a.e. $t \in [0,1]$, $\{\nu\colon\ H(\nu \mid \mu) \leqslant b(t)\}$ is precompact because $H(\cdot \mid \mu)$ is a good rate function. Thus, in particular, $\{g_n(t),\ n \in \mathbb{N}\}$ is precompact. Let $\{\xi_i,\ i \in \mathbb{N}\}$ be a class of continuous bounded convergence-determining functions defined on Σ. For every $i \in \mathbb{N}$ we consider the martingale $(\langle \xi_i, g_n\rangle, \mathcal{F}_n)_{n\in\mathbb{N}}$ defined on the probability space given above, with $\langle\cdot,\cdot\rangle$ as in (11). Since for every $t \in [0,1]$ $g_n(t) \in M^1(\Sigma)$ we have

$$\sup_{n\in\mathbb{N}} \int_0^1 \big\langle \xi_i, g_n(t)\big\rangle\,\mathrm{d}t \leqslant \sup_{x\in\Sigma}\big|\xi_i(x)\big| < \infty,$$

so the real valued sequence $(\langle \xi_i, g_n(t)\rangle)_{n\in\mathbb{N}}$ converges for all i and a.e. t. This and the fact that $\{g_n(t)\}_{n\in\mathbb{N}}$ is precompact for a.e. t imply that $(g_n(t))_{n\in\mathbb{N}}$ is convergent for a.e. t. Hence we can modify g_n on a negligible part of [0,1] in a way that the modified sequence converges in $M^1(\Sigma)$ for all t. We denote by $(\dot\nu_t)_{t\in[0,1]}$ this limit. Let $0 \leqslant j < k \leqslant 2^n$. For every $l \geqslant n$ we have

$$\nu_{\frac{k}{2^n}} - \nu_{\frac{j}{2^n}} = \int_{\frac{j}{2^n}}^{\frac{k}{2^n}} g_l(s)\,\mathrm{d}s.$$

Since $g_l(t) \stackrel{w}{\to} \dot\nu_t$ for a.e. t, it follows from Lebesgue's theorem that for every $f \in C_b(\Sigma)$

$$\lim_{l\to\infty} \int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \big\langle f, g_l(s)\big\rangle\,\mathrm{d}s = \int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \langle f, \dot\nu_s\rangle\,\mathrm{d}s.$$

Furthermore

$$\int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \langle f, \dot{\nu}_s \rangle \,\mathrm{d}s = \left\langle f, \int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \dot{\nu}_s \,\mathrm{d}s \right\rangle,$$

where $\int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \dot{\nu}_s \,\mathrm{d}s$ is interpreted set-wise, i.e., for all $A \in \mathcal{B}_\Sigma$

$$\left(\int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \dot{\nu}_s \,\mathrm{d}s \right)(A) = \int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \dot{\nu}_s(A) \,\mathrm{d}s,$$

and $\int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \langle f, \dot{\nu}_s \rangle \,\mathrm{d}s$ is the limit as $l \to \infty$ of

$$\int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \bigl\langle f, g_l(s) \bigr\rangle \,\mathrm{d}s = \bigl\langle f, \nu_{\frac{k}{2^n}} - \nu_{\frac{j}{2^n}} \bigr\rangle.$$

Hence

$$\nu_{\frac{k}{2^n}} - \nu_{\frac{j}{2^n}} = \int_{\frac{j}{2^n}}^{\frac{k}{2^n}} \dot{\nu}_s \,\mathrm{d}s.$$

Since $(\nu_t - \nu_s)(\Sigma) = t - s$ for every $t \geqslant s \geqslant 0$, $(\nu_t)_{t \in [0,1]}$ is continuous in the variation norm, so we get

$$\nu_t - \nu_s = \int_s^t \dot{\nu}_u \,\mathrm{d}u.$$

Let $\{\eta_i,\ i \in \mathbb{N}\}$ be a dense countable subset of $M^+(\Sigma)$. Since the metric β is derived from a norm (see (18)), $\beta(\cdot, \eta_i)$ is convex for every $i \in \mathbb{N}$, and for a.e. $s \in [0, 1]$

$$\frac{1}{h} \int_s^{s+h} \beta(\dot{\nu}_t, \eta_i) \,\mathrm{d}t \geqslant \beta\left(\frac{1}{h} \int_s^{s+h} \dot{\nu}_t \,\mathrm{d}t, \eta_i \right),$$

then

$$\limsup_{h \to 0} \beta\left(\frac{1}{h} \int_s^{s+h} \dot{\nu}_t \,\mathrm{d}t, \eta_i \right) \leqslant \limsup_{h \to 0} \frac{1}{h} \int_s^{s+h} \beta(\dot{\nu}_t, \eta_i) \,\mathrm{d}t = \beta(\dot{\nu}_s, \eta_i).$$

But we can choose i such that $\beta(\dot{\nu}_s, \eta_i) \leqslant \varepsilon/2$, so

$$\limsup_{h\to 0} \beta\left(\frac{1}{h}\int_s^{s+h} \dot{\nu}_t \,\mathrm{d}t, \dot{\nu}_s\right) \leqslant \limsup_{h\to 0} \frac{1}{h}\int_\Delta^{\Delta+h} \beta(\dot{\nu}_t, \eta_i)\,\mathrm{d}t + \beta(\eta_i, \dot{\nu}_s) \leqslant \varepsilon.$$

Hence $(\nu_t)_{t\in[0,1]}$ admits a weak derivative for a.e. $t \in [0,1]$, and we can conclude.

2. Let $(\nu_t)_{t\in[0,1]} \in \mathcal{AC}_\mu$. For a.e. $s, t \in [0,1]$ such that $s < t$, Jensen's inequality tells us that

$$\begin{aligned}\int_s^t H(\dot{\nu}_u \mid \mu)\,\mathrm{d}u &= (t-s)\int_s^t \frac{1}{t-s} H(\dot{\nu}_u \mid \mu)\,\mathrm{d}u \\ &\geqslant (t-s) H\left(\int_s^t \dot{\nu}_u \frac{\mathrm{d}u}{t-s}\Big|\mu\right) \\ &\geqslant (t-s) H\left(\frac{\nu_t - \nu_s}{t-s}\Big|\mu\right).\end{aligned}$$

Whence $I_\infty(\nu_., \mu) \leqslant \int_0^1 H(\dot{\nu}_u \mid \mu)\,\mathrm{d}u$.

Since for a.e. $u \in [0,1]$ $g_n(u) \xrightarrow{w} \dot{\nu}_u$, we obtain according to Fatou's lemma,

$$I_\infty(\nu_., \mu) \geqslant \liminf_{n\to\infty} \int_0^1 H\big(g_n(u) \mid \mu\big)\,\mathrm{d}u \geqslant \int_0^1 H(\dot{\nu}_u \mid \mu)\,\mathrm{d}u.$$

Thus for all $\nu_. \in \mathcal{AC}_\mu$ $I_\infty(\nu_., \mu) = \int_0^1 H(\dot{\nu}_u \mid \mu)\,\mathrm{d}u$. □

By combining the preceding 3 lemmas we obtain the expected result.

THEOREM 1. – *If $\mu^n \xrightarrow{w} \mu$ then $(l_t^n)_{t\in[0,1]}$ satisfies a LDP on $D[[0,1],(M^+(\Sigma),\beta)]$ with good rate function*

$$I_\infty(\nu_., \mu) = \begin{cases} \int_0^1 H(\dot{\nu}_s \mid \mu)\,\mathrm{d}s & \text{if } \nu_. \in \mathcal{AC}_\mu, \\ \infty & \text{elsewhere.}\end{cases} \tag{21}$$

Proof. – We have $\mathbb{P}^n(\bar{l}_t^n \in C[[0,1],(M^+(\Sigma),\beta)]; \mu^n) = 1$ and for all $\nu_. \notin C[[0,1],(M^+(\Sigma),\beta)]$ $I_\infty(\nu_., \mu) = \infty$. We deduce from Lemma 5 that $(\bar{l}_t^n)_{t\in[0,1]}$ follows a LDP on $C[[0,1],(M^+(\Sigma),\beta)]$ endowed with the topology of pointwise convergence, with good rate function $I_\infty(\nu_., \mu)$. As $(\bar{l}_t^n)_{t\in[0,1]}$ is exponentially tight on $C[[0,1],(M^+(\Sigma),\beta)]$ equipped with the metric β_∞ (Lemma 4), it also satisfies a LDP on $C[[0,1],(M^+(\Sigma),\beta)]$ with the same good rate function. Since $C[[0,1],(M^+(\Sigma),\beta)]$ is closed on $D[[0,1],(M^+(\Sigma),\beta)]$ equipped with β_∞, $(\bar{l}_t^n)_{t\in[0,1]}$ follows a LDP on $D[[0,1],(M^+(\Sigma),\beta)]$ with good rate function $I_\infty(\nu_., \mu)$. Finally, $(\bar{l}_t^n)_{t\in[0,1]}$ and $(l_t^n)_{t\in[0,1]}$ being exponentially equivalent on $D[[0,1],(M^+(\Sigma),\beta)]$ we can conclude, the expression of the rate function resulting from Lemma 6. □

From this LDP we can derive a weak law of large numbers related to microcanonical distributions, as we announced in the introduction.

COROLLARY 2. – *If* $\mu^n \xrightarrow{w} \mu$ *then* $(l_t^n)_{t\in[0,1]}$ *tends in probability to* $(t\mu)_{t\in[0,1]}$ *for the metric* β_∞.

Proof. – Let $\varepsilon > 0$. According to Theorem 1

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}^n\big(\beta_\infty\big(l_t^n, t\mu\big) \geqslant \varepsilon; \mu^n\big) \leqslant - \inf_{B_{\beta_\infty}(t\mu,\varepsilon)^C \cap \mathcal{AC}_\mu} \int_0^1 H(\dot{\nu}_s \mid \mu)\,\mathrm{d}s.$$

But $\int_0^1 H(\dot{\nu}_s \mid \mu)\,\mathrm{d}s = 0$ if and only if $\dot{\nu}_s = \mu$ for a.e. $s \in [0,1]$, i.e., $\nu_s = s\mu$ for a.e. s. Hence $\lim_{n\to\infty} \mathbb{P}^n(\beta_\infty(l_t^n, t\mu) \geqslant \varepsilon; \mu^n) = 0$. □

4. Large deviations for $(L_t^n)_{t\in[0,1]}$

Our aim in this section is to extend the setting of Theorem 1 to general triangular arrays of exchangeable random variables as described in the introduction. The LDP for $(L_t^n)_{t\in[0,1]}$ defined in (1) follows from the fact that, according to Theorem 1, it is a mixture of LDS. Then we can state Theorem 2 by means of a slight modification of Theorem 2.3 in [13].

First we prove that $(L_t^n)_{t\in[0,1]}$ is a mixture of LDS. Let us recall that we denote by $M^{1,n}(\Sigma)$ the subset of $M^1(\Sigma)$ of all atomic measures $\frac{1}{n}\sum_{i=1}^n \delta_{x_i^n}$ for $(x_1^n,\dots,x_n^n) \in \Sigma^n$, possibly with ties, and by Q^n the distribution of $\frac{1}{n}\sum_{i=1}^n \delta_{X_i^n}$ under $\mathbb{P}^n$. Note that Q^n is also the distribution of $\frac{1}{n}\sum_{i=1}^n \delta_{Y_i^n}$ under P^n. We sometimes use the shortcut $f(A) = \inf_{x\in A} f(x)$.

LEMMA 7. – *For all* $n \in \mathbb{N}$ *and all* $\mu \in M^{1,n}(\Sigma)$, $\mathbb{P}^n((x_1^n,\dots,x_n^n) \in \cdot\,; \mu)$ *is a regular version of the distribution of* $(X_1^n,\dots,X_n^n)$ *under* $\mathbb{P}^n$ *conditioned on* $\{\frac{1}{n}\sum_{i=1}^n \delta_{X_i^n} = \mu\}$. *In particular, for all measurable subsets* A *of* $D[[0,1],(M^+(\Sigma),\beta)]$ *we have*

$$\mathbb{P}^n(L_\cdot^n \in A) = \int_{M^{1,n}(\Sigma)} \mathbb{P}^n(l_\cdot^n \in A; \mu)\,Q^n(\mathrm{d}\mu). \tag{22}$$

Proof [From [1], Lemma 5.4]. – Let $\mu \in M^{1,n}(\Sigma)$ and $\rho^n(\mu;\cdot)$ be a regular version of the distribution of $(X_1^n,\dots,X_n^n)$ under $\mathbb{P}^n$ conditioned on $\{\frac{1}{n}\sum_{i=1}^n \delta_{X_i^n} = \mu\}$. Since $(X_1^n,\dots,X_n^n)$ is n-exchangeable, we have for all permutations σ on $\{1,\dots,n\}$

$$\left(X_1^n,\dots,X_n^n, \frac{1}{n}\sum_{i=1}^n \delta_{X_i^n}\right) \stackrel{\mathcal{D}}{=} \left(X_{\sigma(1)}^n,\dots,X_{\sigma(n)}^n, \frac{1}{n}\sum_{i=1}^n \delta_{X_{\sigma(i)}^n}\right).$$

Then $\rho^n(\mu;\cdot)$ is an n-exchangeable measure for almost every $\mu \in M^{1,n}(\Sigma)$. Furthermore, the empirical measure of an n-tuple distributed according to $\rho^n(\mu;\cdot)$ is necessarily $\frac{1}{n}\sum_{i=1}^n \delta_{X_i^n} = \mu$. Hence $\rho^n(\mu;\cdot) \in M^1(\Sigma^n)$ is the distribution of sampling without replacement from an urn which composition is given by μ. Whence $\mathbb{P}^n((x_1^n,\dots,x_n^n) \in \cdot\,; \mu)$ is a regular version of $\rho^n(\mu;\cdot)$. Let A be a measurable subset of $D[[0,1],(M^+(\Sigma),\beta)]$, and $\hat{A}_n$ be the Borel subset of Σ^n defined by $\{L_\cdot^n \in A\} = \{(X_1^n,\dots,X_n^n) \in \hat{A}_n\}$. We have

$$\mathbb{P}^n\big(L^n_\cdot \in A\big) = \mathbb{P}^n\big(\big(X^n_1, \dots, X^n_n\big) \in \hat{A}_n\big) = \int\limits_{M^1(\Sigma)} \rho^n\big(\mu, \hat{A}_n\big) Q^n(\mathrm{d}\mu)$$
$$= \int\limits_{M^1(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in A; \mu\big) Q^n(\mathrm{d}\mu)$$

that is the desired formula. □

The following lemma gives the crucial inequalities in order to prove a LDP for $(L^n_t)_{t\in[0,1]}$.

LEMMA 8. – 1. *Let G be a closed subset of $D[[0,1],(M^+(\Sigma),\beta)]$ and $\mu \in M^1(\Sigma)$ be such that $I_\infty(G,\mu) = \inf_{\nu.\in G} I_\infty(\nu.,\mu) < \infty$. For each $\delta > 0$ there exists a neighborhood U_δ of μ such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log\Big(\sup_{\rho\in U_\delta\cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in G; \rho\big)\Big) \leqslant -I_\infty(G,\mu) + \delta.$$

If $I_\infty(G,\mu) = \infty$, then there exists for each $L \in \mathbb{R}$ a neighborhood U_L of μ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log\Big(\sup_{\rho\in U_L\cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in G; \rho\big)\Big) \leqslant -L.$$

2. *Let O be an open subset of $D[[0,1],(M^+(\Sigma),\beta)]$ and $\mu \in M^1(\Sigma)$ be such that $I_\infty(O,\mu) = \inf_{\nu.\in O} I_\infty(\nu.,\mu) < \infty$. For each $\delta > 0$ there exists a neighborhood U_δ of μ such that*

$$\liminf_{n\to\infty} \frac{1}{n} \log\Big(\inf_{\rho\in U_\delta\cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in O; \rho\big)\Big) \geqslant -I_\infty(O,\mu) - \delta.$$

If $I_\infty(O,\mu) = \infty$, then there exists for each $L \in \mathbb{R}$ a neighborhood U_L of μ such that

$$\liminf_{n\to\infty} \frac{1}{n} \log\Big(\inf_{\rho\in U_L\cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in O; \rho\big)\Big) \geqslant -L.$$

Proof. – We prove the first assertion of the lemma. Suppose for a contradiction that there exist a closed subset G of $D[[0,1],(M^+(\Sigma),\beta)]$ and $\mu \in M^1(\Sigma)$ such that $I_\infty(G,\mu) < \infty$ and there exists a $\delta > 0$ such that for all neighborhoods U of μ

$$\limsup_{n\to\infty} \frac{1}{n} \log\Big(\sup_{\rho\in U\cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in G; \rho\big)\Big) > -I_\infty(G,\mu) + \delta.$$

Hence, for all neighborhoods U of μ there exists a sequence $(n_k)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} n_k = \infty$ and for k large enough

$$\sup_{\rho\in U\cap M^{1,n_k}(\Sigma)} \mathbb{P}^{n_k}\big(l^{n_k}_\cdot \in G; \rho\big) > \exp\big(n_k\big(-I_\infty(G,\mu) + \delta\big)\big).$$

Whence, there exists a sequence $(m_k)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} m_k = \infty$ and for every $k \in \mathbb{N}$

$$\sup_{\rho\in B(\mu,\frac{1}{k})\cap M^{1,m_k}(\Sigma)} \mathbb{P}^{m_k}\big(l_\cdot^{m_k} \in F; \rho\big) > \exp\big(m_k\big(-I_\infty(F,\mu)+\delta\big)\big).$$

For all $k \in \mathbb{N}$ there exists a $\rho_k \in B(\mu, \frac{1}{k})$ such that

$$\begin{aligned}\mathbb{P}^{m_k}\big(l_\cdot^{m_k} \in G; \rho_k\big) &> \sup_{\rho\in B(\mu,\frac{1}{k})\cap M^{1,m_k}(\Sigma)} \big\{\mathbb{P}^{m_k}\big(l_\cdot^{m_k} \in G; \rho\big)\big\} - \exp\big(-m_k^2\big)\\ &> \exp\big(m_k\big(-I_\infty(G,\mu)+\delta\big)\big) - \exp\big(-m_k^2\big).\end{aligned}$$

Hence

$$\limsup_{k\to\infty} \frac{1}{m_k} \log\big(\mathbb{P}^{m_k}\big(l_\cdot^{m_k} \in G; \rho_k\big) + \exp\big(-m_k^2\big)\big) \geqslant -I_\infty(G,\mu)+\delta. \tag{23}$$

But, according to Lemma 1.2.15 in [6] and Theorem 1, we should obtain

$$\begin{aligned}&\limsup_{k\to\infty} \frac{1}{m_k} \log\big(\mathbb{P}^{m_k}\big(l_\cdot^{m_k} \in G; \rho_k\big) + \exp\big(-m_k^2\big)\big)\\ &\quad= \max\left(\limsup_{k\to\infty} \frac{1}{m_k} \log \mathbb{P}^{m_k}\big(l_\cdot^{m_k} \in G; \rho_k\big), \limsup_{k\to\infty} \frac{1}{m_k} \log\exp\big(-m_k^2\big)\right)\\ &\quad= \limsup_{k\to\infty} \frac{1}{m_k} \log \mathbb{P}^{m_k}\big(l_\cdot^{m_k} \in G; \rho_k\big)\\ &\quad\leqslant -I_\infty(G,\mu).\end{aligned}$$

Clearly, the last display cannot hold simultaneously with (23). The proof of the three other inequalities follows the same pattern. □

We recall that $\mathcal{AC}$ is the space of all maps $\nu_t : [0,1] \to M^+(\Sigma)$ such that $\nu_t - \nu_s \in M^+(\Sigma)$ of total mass $t-s$ for all $0 \leqslant s < t$, $\nu_0 = 0$, and which possess a weak derivative for a.e. $t \in [0,1]$ as defined in (5).

THEOREM 2. – *Suppose that L_1^n obeys a LDP on $M^1(\Sigma)$ with good rate function J. Then $(L_t^n)_{t\in[0,1]}$ obeys a LDP on $D[[0,1],(M^+(\Sigma),\beta)]$ with good rate function*

$$I(\nu_\cdot) = I_\infty(\nu_\cdot,\nu_1) + J(\nu_1) = \begin{cases} \int_0^1 H(\dot{\nu_s} \mid \nu_1)\,\mathrm{d}s + J(\nu_1) & \text{if } \nu_\cdot \in \mathcal{AC},\\ \infty & \text{elsewhere.}\end{cases} \tag{24}$$

Proof. – We first prove the upper bound of the LDP. Let G be a closed subset of $D[[0,1],(M^+(\Sigma),\beta)]$, $\varepsilon > 0$ and $L \geqslant 0$. Let $\phi_L^J = \{\nu \in M^1(\Sigma),\ J(\nu) \leqslant L\}$, which is compact since J is a good rate function. Lemma 8 tells us that for every $\mu \in M^1(\Sigma)$ there exists a neighborhood U_μ of μ such that

$$\limsup_{n\to\infty} \frac{1}{n} \log\Big(\sup_{\rho\in U_\mu\cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(l_\cdot^n \in G; \rho\big)\Big) \leqslant -K_\mu + \frac{\varepsilon}{2},$$

where

$$K_\mu = \begin{cases} I_\infty(G,\mu) & \text{if } I_\infty(G,\mu) < \infty, \\ L & \text{otherwise.} \end{cases}$$

Since J is lower semicontinuous U_μ can be modified such that it also satisfies

$$\inf_{\rho \in \bar{U}_\mu} J(\rho) \geqslant J(\mu) - \frac{\varepsilon}{2}.$$

As ϕ_L^J is compact, there exist $\mu_1, \ldots, \mu_k$ such that $\phi_L^J \subset \bigcup_{i=1}^k U_{\mu_i} = C_L$. Hence, there exists an N_0 such that for all $n \geqslant N_0$

$$\begin{aligned} \mathbb{P}^n\big(L^n_\cdot \in G\big) &= \int_{M^1(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in G; \mu\big) Q^n(\mathrm{d}\mu) \\ &\leqslant Q^n\big(C_L^c\big) + \sum_{i=1}^k \int_{U_{\mu_i} \cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in G; \mu\big) Q^n(\mathrm{d}\mu) \\ &\leqslant Q^n\big(C_L^c\big) + \sum_{i=1}^k \exp\left(-n\left(K_{\mu_i} - \frac{\varepsilon}{2}\right)\right) Q^n(U_{\mu_i}). \end{aligned}$$

Whence

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}^n\big(L^n_\cdot \in G\big) &\leqslant \max_{i=1,\ldots,k} \big\{-L, -K_{\mu_i} - J(\mu_i) + \varepsilon\big\} \\ &\leqslant \max_{i=1,\ldots,k} \big\{\big\{-I_\infty(G,\mu_i) - J(\mu_i) + \varepsilon\big\}, -L\big\}. \end{aligned}$$

We obtain the upper bound of the LDP by letting $L \to \infty$ and then $\varepsilon \to 0$.

Now we prove the lower bound of the LDP. Let O be an open subset of $D[[0,1], (M^+(\Sigma), \beta)]$ and $\varepsilon > 0$. Let $\nu_\cdot \in O$ be such that $I(\nu_\cdot) < \infty$. According to Lemma 8 there exists a neighborhood U of ν_1 such that

$$\liminf_{n\to\infty} \frac{1}{n} \log\Big(\inf_{\rho \in U \cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in O; \rho\big)\Big) \geqslant -I_\infty(O, \nu_1) - \varepsilon.$$

Whence

$$\begin{aligned} \mathbb{P}^n\big(L^n_\cdot \in O\big) &\geqslant \int_{U \cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(l^n_\cdot \in O; \rho\big) Q^n(\mathrm{d}\rho) \\ &\geqslant \exp\big(-n\big(I_\infty(O,\nu_1) + \varepsilon\big)\big) Q^n(U). \end{aligned}$$

Then

$$\begin{aligned} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}^n\big(L^n_\cdot \in O\big) &\geqslant -I_\infty(O,\nu_1) - \inf_{\rho \in U} J(\rho) - \varepsilon \\ &\geqslant -I_\infty(\nu_\cdot, \nu_1) - J(\nu_1) - \varepsilon. \end{aligned}$$

We obtain the desired lower bound by letting $\varepsilon \to 0$.

Next we prove that I is a good rate function. We denote by π the projection that maps $D[[0,1],(M^+(\Sigma),\beta)]$ to $M^1(\Sigma)$ by $(\nu_t)_{t\in[0,1]}\mapsto\nu_1$. Suppose for a contradiction that there exists an $\alpha>0$ such that $\phi^I_\alpha=\{\nu_\cdot\in D[[0,1],(M^+(\Sigma),\beta)],\ I(\nu_\cdot)\leqslant\alpha\}$ is not compact. Then there is a sequence $(\nu^n_\cdot)_{n\in\mathbb{N}}\in\phi^I_\alpha\subset C[[0,1],(M^+(\Sigma),\beta)]$ that does not have any convergent subsequence. As $(\nu^n_1)_{n\in\mathbb{N}}\in\phi^J_\alpha$, it admits a convergent subsequence $(\nu^{n_k}_1)_{k\in\mathbb{N}}$ and we put $\lim_{k\to\infty}\nu^{n_k}_1=\eta_1$. Let $(\bar\nu^{n_k}_1)_{k\in\mathbb{N}}$ be such that $\bar\nu^{n_k}_1\in M^{1,n_k}(\Sigma)$ for all $k\in\mathbb{N}$ and $\bar\nu^{n_k}_1\overset{w}{\to}\eta_1$. We have stated in the proof of Theorem 1 that the family $\mathbb{P}^{n_k}(\bar l^{n_k}_\cdot\in\cdot\,;\bar\nu^{n_k}_1)$ follows a LDP on the Polish space $C[[0,1],(M^+(\Sigma),\beta)]$ with a good rate function. Hence it is exponentially tight, i.e. there exists a compact K_{η_1} in $C[[0,1],(M^+(\Sigma),\beta)]$ such that

$$\limsup_{k\to\infty}\frac{1}{n_k}\log\mathbb{P}^{n_k}\left(\bar l^{n_k}_\cdot\in(K_{\eta_1})^c;\bar\nu^{n_k}_1\right)\leqslant-3\alpha.$$

Since $(\nu^{n_k}_\cdot)_{k\in\mathbb{N}}$ has no accumulation point there exists an N_0 such that for all $k\geqslant N_0$ $\nu^{n_k}_\cdot\notin K_{\eta_1}$. As $\{\nu^{n_k}_\cdot,k\geqslant N_0\}$ is closed and $C[[0,1],(M^+(\Sigma),\beta)]$ is metric there are two disjoint open subsets U_D and U_K such that $\{\nu^{n_k}_\cdot,k\geqslant N_0\}\subset U_D$ and $K_{\eta_1}\subset U_K$. The results in Lemmas 7 and 8 are still valid if we replace L^n_t by its linear interpolation $\bar L^n_t$, whence there is a neighborhood V of η_1 such that

$$\begin{aligned}
&\limsup_{k\to\infty}\frac{1}{n_k}\log\mathbb{P}^{n_k}\left(\bar L^{n_k}_\cdot\in\left(U^c_K\cap\pi^{-1}(V)\right)\right)\\
&\quad\leqslant\limsup_{k\to\infty}\frac{1}{n_k}\log\sup_{\gamma\in V\cap M^{1,n_k}(\Sigma)}\mathbb{P}^{n_k}\left(\bar l^{n_k}_\cdot\in U^c_K;\gamma\right)\\
&\quad\leqslant\limsup_{k\to\infty}\frac{1}{n_k}\log\mathbb{P}^{n_k}\left(\bar l^{n_k}_\cdot\in U^c_K;\bar\nu^{n_k}_1\right)+\alpha\\
&\quad\leqslant\limsup_{k\to\infty}\frac{1}{n_k}\log\mathbb{P}^{n_k}\left(\bar l^{n_k}_\cdot\in K^c_{\eta_1};\bar\nu^{n_k}_1\right)+\alpha\leqslant-2\alpha.
\end{aligned}$$

According to the lower bound of the LDP followed by $\bar L^n_\cdot$

$$\begin{aligned}
\liminf_{k\to\infty}\frac{1}{n_k}\log\mathbb{P}^{n_k}\left(\bar L^{n_k}_\cdot\in\left(U_D\cap\pi^{-1}(V)\right)\right)&\geqslant-\inf_{\nu_\cdot\in U_D\cap\pi^{-1}(V)}I(\nu_\cdot)\\
&\geqslant-\alpha.
\end{aligned}$$

But these two inequalities cannot hold simultaneously, hence ϕ^I_α is compact. □

From this LDP we obtain the following weak law of large numbers.

COROLLARY 3. – *If* $\frac{1}{n}\sum_{i=1}^n\delta_{X^n_i}\overset{w}{\to}\mu$ *in* $\mathbb{P}^n$*-probability then* $(L^n_t)_{t\in[0,1]}$ *tends to* $(t\mu)_{t\in[0,1]}$ *in* $\mathbb{P}^n$*-probability for the distance* β_∞.

Proof. – Let $\varepsilon>0$, $F_\varepsilon=B_{\beta_\infty}(t\mu,\varepsilon)^c$, and $\delta>0$ be such that $-I_\infty(F_\varepsilon,\mu)+\delta<0$. According to Lemma 8 there exists a neighborhood U_δ of μ such that

$$\limsup_{n\to\infty}\frac{1}{n}\log\Big(\sup_{\rho\in U_\delta\cap M^{1,n}(\Sigma)}\mathbb{P}^n\left(l^n_\cdot\in F_\varepsilon;\rho\right)\Big)\leqslant-I_\infty(F_\varepsilon,\mu)+\delta.$$

Let $\eta > 0$ be such that $B_\beta(\mu, \eta) \subset U_\delta$. We have

$$\begin{aligned}\mathbb{P}^n\big(\beta_\infty\big(L^n_\cdot, t\mu\big) \geqslant \varepsilon\big) &= \mathbb{P}^n\big(\beta_\infty\big(L^n_\cdot, t\mu\big) \geqslant \varepsilon, \beta\big(L^n_1, \mu\big) \geqslant \eta\big)\\ &\quad + \mathbb{P}^n\big(\beta_\infty\big(L^n_\cdot, t\mu\big) \geqslant \varepsilon, \beta\big(L^n_1, \mu\big) < \eta\big).\end{aligned}$$

Since $\lim_{n\to\infty} \mathbb{P}^n(\beta(L^n_1, \mu) \geqslant \eta) = 0$ we obtain

$$\lim_{n\to\infty} \mathbb{P}^n\big(\beta_\infty\big(L^n_\cdot, t\mu\big) \geqslant \varepsilon, \beta\big(L^n_1, \mu\big) \geqslant \eta\big) = 0.$$

By virtue of Lemma 8

$$\begin{aligned}\mathbb{P}^n\big(\beta_\infty\big(L^n_\cdot, t\mu\big) \geqslant \varepsilon, \beta\big(L^n_1, \mu\big) < \eta\big) &= \int_{M^1(\Sigma)} \mathbb{P}^n\big(\beta_\infty\big(l^n_\cdot, t\mu\big) \geqslant \varepsilon, \beta(\rho, \mu) < \eta; \rho\big) Q^n(\mathrm{d}\rho)\\ &= \int_{B_\beta(\mu,\eta)} \mathbb{P}^n\big(\beta_\infty\big(l^n_\cdot, t\mu\big) \geqslant \varepsilon; \rho\big) Q^n(\mathrm{d}\rho)\\ &\leqslant \sup_{\rho \in B_\beta(\mu,\eta)\cap M^{1,n}(\Sigma)} \mathbb{P}^n\big(\beta_\infty\big(l^n_\cdot, t\mu\big) \geqslant \varepsilon; \rho\big)\\ &\leqslant \exp n\big(-I_\infty(F_\varepsilon, \mu) + \delta\big),\end{aligned}$$

so $\lim_{n\to\infty} \mathbb{P}^n(\beta_\infty(L^n_\cdot, t\mu) \geqslant \varepsilon) = 0$. □

5. Applications

In this section we consider several applications of Theorem 2.

5.1. The Curie–Weiss model

The Curie–Weiss model is a well known toy model of statistical mechanics. Let $\Sigma = \{-1, 1\}$ and λ_p be the Bernoulli measure on Σ with parameter p ($p \in]0, 1[$). For every $n \in \mathbb{N}$, we associate to each configuration $(x^n_1, \ldots, x^n_n) \in \Sigma^n$ of the system the Hamiltonian

$$\begin{aligned}H_n\big(x^n_1, \ldots, x^n_n\big) &= ng\left(\sum_{i=1}^n \frac{x^n_i}{n}\right)\\ &= n\left(\frac{J_0}{2}\left(\sum_{i=1}^n \frac{x^n_i}{n}\right)^2 + h\left(\sum_{i=1}^n \frac{x^n_i}{n}\right)\right),\end{aligned}$$

where J_0 and h are constants representing a ferro-magnetic coupling and an external magnetic field respectively. The Hamiltonian H_n is in fact a functional of the quantity $\frac{1}{n}\sum_{i=1}^n x^n_i$ called the total magnetization of the system.

In the setting of equilibrium statistical mechanics two joint probability distributions appear to be significant. The first one is the *microcanonical ensemble* which is obtained by conditioning the distribution $\lambda_p^{\otimes n}$ on the energy shell

$$A^{u,n} = \big\{\big(x^n_1, \ldots, x^n_1\big) \in \Sigma^n\colon\ H_n\big(x^n_1, \ldots, x^n_n\big) = u\big\}$$

where $u \in \mathbb{R}$. In general cases, in order to avoid problems with the existence of regular conditioned probabilities, $\lambda_p^{\otimes n}$ is conditioned on the thickened energy shell

$$A^{u,n,r} = \{(x_1^n, \ldots, x_1^n) \in \Sigma^n\colon\ H_n(x_1^n, \ldots, x_n^n) \in [u-r, u+r]\}$$

with $r > 0$. In the case we are interested in, conditioning on the event

$$B^{u,n} = \left\{(x_1^n, \ldots, x_1^n) \in \Sigma^n\colon\ \frac{1}{n}\sum_{i=1}^n \delta_{x_i^n} \in \{\mu_1^n, \mu_2^n\}\right\}$$

seems to be more accurate. Here, the $\mu_i^n \in M^{1,n}(\Sigma)$ are solutions of $ng(\int_\Sigma x\mu(\mathrm{d}x)) = u_n$, u_n being the closest element to u in the set $\{ng(\int_\Sigma x\mu(\mathrm{d}x)),\ \mu \in M^{1,n}(\Sigma)\}$. There are at most two measures solutions of this problem. Thus, the microcanonical ensemble is an equally-likely mixture of the probabilities $\mathbb{P}^n(\cdot\,;\mu_i^n)$ associated to sampling without replacement in the "urn" μ_i^n. Our study allows us to give the LDP for the empirical measure process $(l_t^n)_{t\in[0,1]}$ under the microcanonical ensemble. Indeed, $\mu_i^n \xrightarrow{w} \lambda_{1/2}$, so according to Theorem 1 and Theorems 2.1 and 2.2 in [9] the distribution of $(l_t^n)_{t\in[0,1]}$ under the microcanonical distribution follows a LDP with good rate function

$$I_\infty(\nu_\cdot, \lambda_{1/2}) = \begin{cases} \int_0^1 H(\dot{\nu}_s \mid \lambda_{1/2})\,\mathrm{d}s & \text{if } \nu_\cdot \in \mathcal{AC}_\mu, \\ \infty & \text{elsewhere.} \end{cases} \tag{25}$$

The second probability measure that appears in the study of equilibrium is the *canonical ensemble,* defined for all subsets B of Σ^n by

$$P_{n,\beta}(B) = \int_B \frac{\exp(-\beta H_n(x_1^n, \ldots, x_n^n))}{Z_n(\beta)} \prod_{i=1}^n \lambda_p(x_i^n),$$

where $Z_n(\beta)$ stands for the normalization constant

$$Z_n(\beta) = \int_{\Sigma^n} \exp(-\beta H_n(x_1^n, \ldots, x_n^n)) \prod_{i=1}^n \lambda_p(x_i^n).$$

The coordinate maps $(X_1^n, \ldots, X_n^n)$ on Σ^n distributed according to $P_{n,\beta}$ are n-exchangeable random variables. The LDP for the distribution of $\frac{1}{n}\sum_{i=1}^n X_i^n$ under $P_{n,\beta}$ has been done by Ellis [11]. Orey gives in [15] the LDP satisfied by the distribution of the empirical field under the canonical ensemble. Our study allows us to give the LDP for the empirical measure process $(L_t^n)_{t\in[0,1]}$, under the probability $P_{n,\beta}$. This LDP allows to consider applications involving randomly selected segments of the n-tuple $(X_1^n, \ldots, X_n^n)$, having a data dependent location and length. Now we look for this LDP. We know that the distribution of $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i^n$ under $P_{n,\beta}$ obeys a LDP on $[-1,1]$ with good rate function

$$I(z) = I_{\mathcal{C}}(z) - \beta g(z) - \inf_{z\in[-1,1]}[I_{\mathcal{C}}(z) - \beta g(z)],$$

where $I_{\mathcal{C}}$ is the rate function of Cramer's theorem for Bernoulli random variables (see [11]). Since $L_1^n(1)$ and $\bar{X}_n$ are one-to-one linked by $\bar{X}_n = 2L_1^n(1) - 1$, L_1^n follows a LDP on $M^1(\Sigma)$ with good rate function $J(\nu_1) = I(\int_\Sigma x\nu_1(\mathrm{d}x))$. But, for every $\nu_1 \in M^1(\Sigma)$ $I_{\mathcal{C}}(\int_\Sigma x\nu_1(\mathrm{d}x)) = H(\nu_1|\lambda_p)$. Hence

$$J(\nu_1) = H(\nu_1 \mid \lambda_p) - \beta g\biggl(\int_\Sigma x\nu_1(\mathrm{d}x)\biggr) - \inf_{\nu_1 \in M^1(\Sigma)}\biggl[H(\nu_1 \mid \lambda_p) - \beta g\biggl(\int_\Sigma x\nu_1(\mathrm{d}x)\biggr)\biggr].$$

Whence, according to Theorem 2, the distribution of $(L_t^n)_{t\in[0,1]}$ under the canonical ensemble follows a LDP on $D[[0,1],(M^+(\Sigma),\beta)]$ with good rate function

$$I_\beta(\nu.) = \begin{cases} \int_0^1 H(\dot{\nu}_s \mid \nu_1)\,\mathrm{d}s + H(\nu_1 \mid \lambda_p) - \beta g\bigl(\int_\Sigma x\nu_1(\mathrm{d}x)\bigr) - C & \text{if } \nu. \in \mathcal{AC}, \\ \infty & \text{otherwise,} \end{cases}$$

where $C = \inf_{\nu_1 \in M^1(\Sigma)}[H(\nu_1 \mid \lambda_p) - \beta g(\int_\Sigma x\nu_1(\mathrm{d}x))]$. The following result helps us in simplifying the expression of I.

LEMMA 9. – *For every $\nu. \in \mathcal{AC}$ and every $\lambda \in M^1(\Sigma)$*

$$\int_0^1 H(\dot{\nu}_s \mid \nu_1)\,\mathrm{d}s + H(\nu_1 \mid \lambda) = \int_0^1 H(\dot{\nu}_s \mid \lambda)\,\mathrm{d}s. \tag{26}$$

Proof. – Let $\nu. \in \mathcal{AC}$ and $\lambda \in M^1(\Sigma)$. First we suppose that ν_1 and λ are such that $H(\nu_1 \mid \lambda) = \infty$. Hence, according to Jensen's inequality

$$\int_0^1 H(\dot{\nu}_s \mid \lambda)\,\mathrm{d}s \geqslant H\biggl(\int_0^1 \dot{\nu}_s\,\mathrm{d}s \biggm| \lambda\biggr) \geqslant H(\nu_1 \mid \lambda) = \infty,$$

so in this case $\int_0^1 H(\dot{\nu}_s \mid \nu_1)\,\mathrm{d}s + H(\nu_1 \mid \lambda) = \int_0^1 H(\dot{\nu}_s \mid \lambda)\,\mathrm{d}s$.

Suppose now that $H(\nu_1|\lambda) < \infty$. Since for all $A \in \mathcal{B}_\Sigma$ $t \mapsto \nu_t(A)$ is an increasing map, $\nu_1(A) = 0$ implies that $\dot{\nu}_s(A) = 0$ for every $s \in [0,1]$. Hence $\dot{\nu}_s$ is absolutely continuous w.r.t. ν_1, and we obtain for every $s \in [0,1]$

$$\begin{aligned} H(\dot{\nu}_s \mid \nu_1) + H(\nu_1 \mid \lambda) &= \int_\Sigma \log\frac{\mathrm{d}\dot{\nu}_s}{\mathrm{d}\nu_1}\,\mathrm{d}\dot{\nu}_s + \int_\Sigma \log\frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda}\,\mathrm{d}\nu_1 \\ &= \int_\Sigma \log\frac{\mathrm{d}\dot{\nu}_s}{\mathrm{d}\nu_1}\,\mathrm{d}\dot{\nu}_s + \int_\Sigma \log\frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda}\,\mathrm{d}\dot{\nu}_s + \int_\Sigma \log\frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda}\,\mathrm{d}\nu_1 - \int_\Sigma \log\frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda}\,\mathrm{d}\dot{\nu}_s \\ &= \int_\Sigma \log\frac{\mathrm{d}\dot{\nu}_s}{\mathrm{d}\lambda}\,\mathrm{d}\dot{\nu}_s + \int_\Sigma \log\frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda}\,\mathrm{d}\nu_1 - \int_\Sigma \log\frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda}\,\mathrm{d}\dot{\nu}_s \\ &= H(\dot{\nu}_s \mid \lambda) + \int_\Sigma \log\frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda}\,\mathrm{d}\nu_1 - \int_\Sigma \log\frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda}\,\mathrm{d}\dot{\nu}_s. \end{aligned}$$

Let us denote by f the measurable map $f = \log \frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda}$. To complete the proof it is sufficient to show that

$$\int_0^1 \left(\int_\Sigma f \, \mathrm{d}\dot{\nu}_s \right) \mathrm{d}s = \int_\Sigma f \, \mathrm{d}\nu_1.$$

For step functions $f = \sum_{i=1}^k \alpha_i 1_{A_i}$ this relation follows from the definition of $\dot{\nu}_s$. For general f's we let f_+ be the positive part of f and we denote by $(f_n)_{n\in N}$ an increasing sequence of step functions that converges to f_+. We obtain

$$\int_0^1 \left(\int_\Sigma f_+ \, \mathrm{d}\dot{\nu}_s \right) \mathrm{d}s = \int_0^1 \left(\lim_{n\to\infty} \int_\Sigma f_n \, \mathrm{d}\dot{\nu}_s \right) \mathrm{d}s \geqslant \int_0^1 \left(\int_\Sigma f_n \, \mathrm{d}\dot{\nu}_s \right) \mathrm{d}s \geqslant \int_\Sigma f_n \, \mathrm{d}\nu_1,$$

so $\int_0^1 (\int_\Sigma f_+ \, \mathrm{d}\dot{\nu}_s) \, \mathrm{d}s \geqslant \int_\Sigma f_+ \, \mathrm{d}\nu_1$, and according to Fatou's lemma

$$\begin{aligned}
\int_\Sigma f_+ \, \mathrm{d}\nu_1 \geqslant \int_\Sigma f_n \, \mathrm{d}\nu_1 &= \int_0^1 \left(\int_\Sigma f_n \, \mathrm{d}\dot{\nu}_s \right) \mathrm{d}s \\
&\geqslant \liminf_{n\to\infty} \int_0^1 \left(\int_\Sigma f_n \, \mathrm{d}\dot{\nu}_s \right) \mathrm{d}s \\
&\geqslant \int_0^1 \left(\int_\Sigma \liminf_{n\to\infty} f_n \, \mathrm{d}\dot{\nu}_s \right) \mathrm{d}s = \int_0^1 \left(\int_\Sigma f \, \mathrm{d}\dot{\nu}_s \right) \mathrm{d}s.
\end{aligned}$$

Whence $\int_0^1 (\int_\Sigma f_+ \, \mathrm{d}\dot{\nu}_s) \, \mathrm{d}s = \int_\Sigma f_+ \, \mathrm{d}\nu_1 < \infty$, since $H(\nu_1 \mid \lambda) < \infty$. So it follows that

$$\int_0^1 \left(\int_\Sigma \log \frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda} \, \mathrm{d}\dot{\nu}_s \right) \mathrm{d}s = \int_\Sigma \log \frac{\mathrm{d}\nu_1}{\mathrm{d}\lambda} \, \mathrm{d}\nu_1,$$

and this ends the proof. □

By virtue of Lemma 9 we obtain

$$I_\beta(\nu.) = \begin{cases} \int_0^1 H(\dot{\nu}_s \mid \lambda_p) \, \mathrm{d}s - \beta g\left(\int_\Sigma x \nu_1(\mathrm{d}x) \right) - C & \text{if } \nu. \in \mathcal{AC}, \\ \infty & \text{elsewhere.} \end{cases}$$

We can prove that $(L_t^n)_{t\in[0,1]}$ follows a LDP in this set-up another way. According to Theorem 1 in [4] the distribution of $(L_t^n)_{t\in[0,1]}$ under $\lambda_p^{\otimes n}$ (i.e., $X_1^n, \ldots, X_n^n$ being independent and identically distributed according to λ_p) follows a LDP with good rate function

$$\tilde{I}(\nu., \lambda_p) = \begin{cases} \int_0^1 H(\dot{\nu}_s \mid \lambda_p) \, \mathrm{d}s & \text{if } \nu. \in \mathcal{AC}, \\ \infty & \text{elsewhere.} \end{cases}$$

Hence, from Varadhan's lemma we know that the distribution of $(L_t^n)_{t\in[0,1]}$ under $P_{n,\beta}$ follows a LDP with good rate function

$$\bar{I}_\beta(\nu.) = \begin{cases} \int_0^1 H(\dot{\nu}_s \mid \lambda_p)\,\mathrm{d}s - \beta g\left(\int_\Sigma x\nu_1(\mathrm{d}x)\right) - \bar{C} & \text{if } \nu. \in \mathcal{AC}, \\ \infty & \text{otherwise,} \end{cases}$$

where

$$\bar{C} = \inf_{\nu.\in D[[0,1],(M^+(\Sigma),\beta)]} \left[\int_0^1 H(\dot{\nu}_s \mid \lambda_p)\,\mathrm{d}s - \beta g\left(\int_\Sigma x\nu_1(\mathrm{d}x)\right)\right].$$

It is sufficient, in order to prove the equality of the rate functions, to prove that $C = \bar{C}$. We have

$$\begin{aligned}
\bar{C} &= \inf_{\nu.\in D[[0,1],(M^+(\Sigma),\beta)]} \left[\int_0^1 H(\dot{\nu}_s \mid \lambda_p)\,\mathrm{d}s - \beta g\left(\int_\Sigma x\nu_1(\mathrm{d}x)\right)\right] \\
&= \inf_{\mu\in M^1(\Sigma)} \left\{ \inf_{\nu.:\ \nu_1=\mu} \left(\int_0^1 H(\dot{\nu}_s \mid \mu)\,\mathrm{d}s + H(\mu \mid \lambda_p) - \beta g\left(\int_\Sigma x\mu(\mathrm{d}x)\right)\right)\right\} \\
&= \inf_{\mu\in M^1(\Sigma)} \left\{ H(\mu \mid \lambda_p) - \beta g\left(\int_\Sigma x\mu(\mathrm{d}x)\right)\right\} = C,
\end{aligned}$$

the equality of I and $\bar{I}$ follows.

5.2. Infinite exchangeable random variables

Let $(X_1,\ldots,X_n,\ldots)$ be an infinite exchangeable sequence of Σ-valued random variables defined on a probability space $(\Omega,\mathcal{A},\mathbb{P})$. For all $n\in\mathbb{N}$ $(X_1^n,\ldots,X_n^n) = (X_1,\ldots,X_n)$ is an n-exchangeable random vector, and according to de Finetti's theorem for any Borel subset of Σ^n

$$\mathbb{P}\big((X_1,\ldots,X_n)\in A\big) = \int_\Theta P_\theta\big((X_1,\ldots,X_n)\in A\big)\gamma(\mathrm{d}\theta),$$

where γ is a probability measure on a closed subset Θ of $M^1(\Sigma)$, and for every $\theta\in\Theta$, P_θ is a probability measure defined on $(\Omega,\mathcal{A})$ such that $X_1,\ldots,X_n,\ldots$ are independent and identically distributed under P_θ. From [9] we know that provided Θ is compact,

$$L_1^n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i} = \frac{1}{n}\sum_{i=1}^n \delta_{X_i^n}$$

follows a LDP on $M^1(\Sigma)$ with good rate function $J(\nu_1) = \inf_{\theta\in\Theta} H(\nu_1 \mid \pi_\theta)$, where $\pi_\theta = P_\theta \circ X_1^{-1}$. Hence, according to Theorem 2, $(L_t^n)_{t\in[0,1]}$ follows a LDP on

$D[[0,1],(M^+(\Sigma),\beta)]$ with good rate function

$$I(\nu_\cdot)=\begin{cases}\int_0^1 H(\dot\nu_s\mid\nu_1)\,\mathrm{d}s+\inf_{\theta\in\Theta}H(\nu_1\mid\pi_\theta) & \text{if } \nu_\cdot\in\mathcal{AC},\\ \infty & \text{elsewhere.}\end{cases}$$

Now, we give a direct proof (without Theorem 2) of this result. Since the mapping from Σ^n to $D[[0,1],(M^+(\Sigma),\beta)]$ defined by L^n_t is continuous, it is an immediate consequence of de Finetti's theorem that for any measurable subset A of $D[[0,1],(M^+(\Sigma),\beta)]$

$$\mathbb{P}\big(L^n_\cdot\in A\big)=\int_\Theta P_\theta\big(L^n_\cdot\in A\big)\gamma(\mathrm{d}\theta).$$

Hence, according to Theorems 2.1, 2.2 in [9], it is sufficient to prove that the family $(P^n_\theta=P_\theta\circ(L^n_\cdot)^{-1},\theta\in\Theta)$ is exponentially continuous to establish the LDP for the distribution of $(L^n_t)_{t\in[0,1]}$ under $\mathbb{P}$. In other words we have to prove that for any converging sequence $\theta^n\xrightarrow{w}\theta$ in Θ and any measurable subset A of $D[[0,1],(M^+(\Sigma),\beta)]$

$$-\inf_{\nu_\cdot\in A^o}\tilde I(\nu_\cdot,\pi_\theta)\leqslant\liminf_{n\to\infty}\frac{1}{n}\log P^n_{\theta^n}(A)\leqslant\limsup_{n\to\infty}\frac{1}{n}\log P^n_{\theta^n}(A)\leqslant-\inf_{\nu_\cdot\in\bar A}\tilde I(\nu_\cdot,\pi_\theta)$$

where $\tilde I(\nu_\cdot,\pi_\theta)$ is defined on $D[[0,1],(M^+(\Sigma),\beta)]$ by

$$\tilde I(\nu_\cdot,\pi_\theta)=\begin{cases}\int_0^1 H(\dot\nu_s\mid\pi_\theta)\,\mathrm{d}s & \text{if } \nu_\cdot\in\mathcal{AC},\\ \infty & \text{elsewhere.}\end{cases}$$

Let $(t_0=0<t_1,\dots,t_{d-1}<t_d\leqslant 1)$ be a strictly ordered $(d+1)$-tuple. We first look for the LDP satisfied by the distribution of $(L^n_{t_0},\dots,L^n_{t_d})$ under P_{θ^n}. Since $X_1,\dots,X_n$ are independent under P_{θ^n} the random empirical measures $L^n_{t_1}-L^n_{t_0},L^n_{t_2}-L^n_{t_1},\dots,L^n_{t_d}-L^n_{t_{d-1}}$ are also independent. It follows from [2] that the distribution of each $L^n_{t_i}-L^n_{t_{i-1}}$ $(1\leqslant i\leqslant d)$ under P_{θ^n} satisfies a LDP on $M^+(\Sigma)$ with good rate function

$$I_i(\nu_i)=(t_i-t_{i-1})H\left(\frac{\nu_i}{t_i-t_{i-1}}\middle|\pi_\theta\right).$$

We deduce from Lemmas 2.7, 2.8 in Lynch and Sethuraman [14] that $(L^n_{t_1}-L^n_{t_0},L^n_{t_2}-L^n_{t_1},\dots,L^n_{t_d}-L^n_{t_{d-1}})$ satisfies a LDP on $M^+(\Sigma)^d$ with good rate function

$$I_{(t_0,\dots,t_d)}(\nu_1,\dots,\nu_d)=\sum_{i=1}^d(t_i-t_{i-1})H\left(\frac{\nu_i}{t_i-t_{i-1}}\middle|\pi_\theta\right).$$

Finally, we deduce from Theorem 1 in [4] that the distribution of $(L^n_t)_{t\in[0,1]}$ under P_{θ^n} follows a LDP with good rate function $\tilde I(\nu_\cdot,\pi_\theta)$. Whence the family $(P^n_\theta,\theta\in\Theta)$ is exponentially continuous, and we deduce from Theorems 2.1, 2.2 in [9] that the distribution of $(L_t^n)_{t\in[0,1]}$ under $\mathbb{P}$ follows a LDP with good rate function

$$\bar{I}(\nu_\cdot)=\begin{cases}\inf_{\theta\in\Theta}\int_0^1 H(\dot{\nu}_s\mid\pi_\theta)\,\mathrm{d}s & \text{if } \nu_\cdot\in\mathcal{AC},\\ \infty & \text{elsewhere.}\end{cases}$$

Next we show that the rate functions I and $\bar{I}$ are equals. From Lemma 9 we know that for all $\theta\in\Theta$ and for all $\nu_\cdot\in\mathcal{AC}$

$$\begin{aligned}\int_0^1 H(\dot{\nu}_s\mid\pi_\theta)\,\mathrm{d}s &= \int_0^1 H(\dot{\nu}_s\mid\nu_1)\,\mathrm{d}s + H(\nu_1\mid\pi_\theta)\\ &\geqslant \int_0^1 H(\dot{\nu}_s\mid\nu_1)\,\mathrm{d}s + \inf_{\theta\in\Theta} H(\nu_1\mid\pi_\theta).\end{aligned}$$

Hence $\bar{I}\geqslant I$. For all $\nu_\cdot\in\mathcal{AC}$ and all $\varepsilon>0$ there exists an $\alpha\in\Theta$ such that $H(\nu_1\mid\pi_\alpha)\leqslant \inf_{\theta\in\Theta}H(\nu_1\mid\pi_\theta)+\varepsilon$, hence

$$\begin{aligned}\int_0^1 H(\dot{\nu}_s\mid\nu_1)\,\mathrm{d}s + H(\nu_1\mid\pi_\alpha) &\leqslant \int_0^1 H(\dot{\nu}_s\mid\nu_1)\,\mathrm{d}s + \inf_{\theta\in\Theta} H(\nu_1\mid\pi_\theta)+\varepsilon,\\ \int_0^1 H(\dot{\nu}_s\mid\pi_\alpha)\,\mathrm{d}s &\leqslant \int_0^1 H(\dot{\nu}_s\mid\nu_1)\,\mathrm{d}s + \inf_{\theta\in\Theta} H(\nu_1\mid\pi_\theta)+\varepsilon,\\ \inf_{\theta\in\Theta}\int_0^1 H(\dot{\nu}_s\mid\pi_\theta)\,\mathrm{d}s &\leqslant \int_0^1 H(\dot{\nu}_s\mid\nu_1)\,\mathrm{d}s + \inf_{\theta\in\Theta} H(\nu_1\mid\pi_\theta)+\varepsilon.\end{aligned}$$

We obtain $I\geqslant\bar{I}$ by letting $\varepsilon\to 0$.

5.3. Sampling with and without replacement

Let $((X_i^n)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ be a triangular array of Σ-valued random variables such that for every $n\in\mathbb{N}$ $X_1^n,\ldots,X_n^n$ are independent and identically distributed according to $\mu^n\in M^1(\Sigma)$. We suppose that $\mu^n\xrightarrow{w}\mu\in M^1(\Sigma)$. From [2] we know that L_1^n obeys a LDP on $M^1(\Sigma)$ with good rate function

$$J(\nu_1)=H(\nu_1\mid\mu).$$

Hence, according to Theorem 2 and Lemma 9 $(L_t^n)_{t\in[0,1]}$ obeys a LDP on $D[[0,1],(M^+(\Sigma),\beta)]$ with good rate function

$$I(\nu_\cdot,\mu)=\begin{cases}\int_0^1 H(\dot{\nu}_s\mid\mu)\,\mathrm{d}s & \text{if } \nu_\cdot\in\mathcal{AC},\\ \infty & \text{elsewhere.}\end{cases}$$

This set-up obviously includes the case where $X_1^n, \ldots, X_n^n$ are given by sampling *with* replacement in an urn whose composition is given by $\mu^n = \frac{1}{n}\sum_{i=1}^n \delta_{y_i^n}$. Let us recall that according to Theorem 1 the rate function of the LDP associated to sampling *without* replacement in the same urn and under the same constraint $\mu^n \xrightarrow{w} \mu$ is

$$I_\infty(\nu_\cdot, \mu) = \begin{cases} \int_0^1 H(\dot{\nu}_s \mid \mu)\,\mathrm{d}s & \text{if } \nu_\cdot \in \mathcal{AC}_\mu, \\ \infty & \text{elsewhere,} \end{cases}$$

i.e., the rate function of the sampling *with* replacement case relativized to μ.

5.4. Random permutations of random processes

Let $(Y_1, Y_2, \ldots, Y_n, \ldots)$ be a Σ-valued process satisfying a Sanov result, and let $((X_i^n)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ be a finite exchangeable triangular array of random variables defined as follows: For every $n \in \mathbb{N}$ we uniformly choose a random permutation σ^n on $\{1, \ldots, n\}$ and we put $X_i^n = Y_{\sigma^n(i)}$. The resulting process describes the transmission of the random signal Y^n chopped into n pieces of equal length $(Y_1, Y_2, \ldots, Y_n)$, each piece being transmitted to the same destination by different paths. The order of arrival of the pieces (given by σ^n) is assumed to be uniform and independent of Y^n. We consider here the particular case where the spring process is a Markov chain. Let $(Y_1, Y_2, \ldots, Y_n, \ldots)$ be a Σ-valued Markov chain with probability transition $p(x, \mathrm{d}y)$. We suppose that $p(x, \mathrm{d}y)$ satisfies the Feller property, i.e., for all $f \in C_b(\Sigma)$ the function

$$x \in \Sigma \mapsto (pf)(x) = \int_\Sigma f(y)p(x, \mathrm{d}y)$$

is continuous. It is also assumed that there exist integers $0 < l \leqslant N$ and a constant $M \geqslant 1$ such that for all $x, x' \in \Sigma$

$$p^l(x, \cdot) \leqslant \frac{M}{N}\sum_{m=1}^N p^m(x', \cdot),$$

where $p^m(x, \cdot)$ is the m-step transition probability for initial condition y, given by

$$p^{m+1}(x, \cdot) = \int_\Sigma p^m(y, \cdot)(x, \mathrm{d}y).$$

We know that for any starting point $L_1^n = \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}$ satisfies a LDP with good rate function

$$J(\nu_1) = \sup_{u\in\mathcal{U}(\Sigma)}\left\{\int_\Sigma \log\left(\frac{u}{pu}\right)\mathrm{d}\nu_1\right\},$$

where $\mathcal{U}(\Sigma)$ denotes the set of $u \in C_b(\Sigma)$ satisfying $u \geqslant 1$ on Σ (see [7]). We let $((X_i^n)_{1\leqslant i\leqslant n})_{n\in\mathbb{N}}$ be defined as above. According to Theorem 2 $(L_t^n)_{t\in[0,1]}$ follows a LDP on $D[[0,1],(M^+(\Sigma),\beta)]$ with good rate function

$$I(\nu_\cdot)=\begin{cases}\int_0^1 H(\dot{\nu}_s \mid \nu_1)\,\mathrm{d}s+\sup_{u\in\mathcal{U}(\Sigma)}\left\{\int_\Sigma \log\left(\frac{u}{pu}\right)\mathrm{d}\nu_1\right\} & \text{if } \nu_\cdot\in\mathcal{AC},\\ \infty & \text{elsewhere.}\end{cases}$$

Acknowledgements

This work is part of my PhD thesis. I would like to thank my supervisor Professor Francis Comets for a wealth of advice and encouragement.

REFERENCES

[1] Aldous D.J., Exchangeability and related topics, in: Ecole d'été de probabilités de Saint-Flour XIII-1983, Lecture Notes in Math., Vol. 1117, Springer, Berlin, 1985.
[2] Baxter J.R., Jain N.C., Comparison principle for large deviations, Proc. Amer. Math. Soc. 103 (1988) 1235–1240.
[3] Dawson D.A., Gartner J., Multilevel large deviations and interacting diffusions, Probab. Theory Related Fields 98 (1994) 423–487.
[4] Dembo A., Zajic T., Large deviations: From empirical mean and measure to partial sums process, Stoch. Proc. and Appl. 57 (1995) 191–224.
[5] Dembo A., Zeitouni O., Large deviations of sub-sampling from individual sequences, Statist. and Probab. Lett. 27 (1996) 201–205.
[6] Dembo A., Zeitouni O., Large Deviations Techniques and Applications, Springer, New York, 1998.
[7] Deuschel J.D., Stroock D.W., Large Deviations, Academic Press, Boston, 1989.
[8] Diaconis P., Freedman D., Finite exchangeable sequences, Ann. Probab. 8 (1980) 745–764.
[9] Dinwoodie I.H., Zabell S.L., Large deviations for exchangeable random vectors, Ann. Probab. 20 (1992) 1147–1166.
[10] Dupuis P., Ellis R.S., A Weak Convergence Approach to the Theory of Large Deviations, Wiley, New York, 1997.
[11] Ellis R.S., Entropy, Large Deviations and Statistical Mechanics, Springer, New York, 1985.
[12] Feller W., An Introduction to Probability Theory and its Applications, Vol. 2, Wiley, New York, 1971.
[13] Grunwald M., Sanov results for Glauber spin-glass dynamics, Probab. Theory Related Fields 106 (1996) 187–232.
[14] Lynch J., Sethuraman J., Large deviations for processes with independent increments, Ann. Probab. 15 (1987) 610–627.
[15] Orey S., Large deviations for the empirical field of Curie–Weiss models, Stochastics 25 (1988) 3–14.
[16] Papangelou F., On the gaussian fluctuations of the critical Curie–Weiss model in statistical mechanics, Probab. Theory Related Fields 83 (1989) 265–278.
[17] Papangelou F., Large deviations and the internal fluctuations of critical mean-field systems, Stoch. Proc. and Appl. 36 (1990) 1–14.
[18] Stroock D.W., Zeitouni O., Microcanonical distributions, Gibbs states, and the equivalence of ensembles, in: Durrettt R., Kesten H. (Eds.), Festschrift in Honour of F. Spitzer, Birkhäuser, Basel, 1991, pp. 399–424.