Electronic Journal of Probability

Vol. 7 (2002) Paper No. 4, pages 1–32.

Journal URL
http://www.math.washington.edu/~ejpecp/
Paper URL
http://www.math.washington.edu/~ejpecp/EjpVol7/paper4.abs.html

A CRAMÉR TYPE THEOREM FOR WEIGHTED RANDOM VARIABLES

Jamal Najim
Laboratoire Modal-X, Université Paris 10- Nanterre
200, Av. de la République, 92001 Nanterre cedex, France
najim@tsi.enst.fr

Abstract A Large Deviation Principle (LDP) is proved for the family $\frac{1}{n}\sum_1^n f(x_i^n)\cdot Z_i$ where $\frac{1}{n}\sum_1^n \delta_{x_i^n}$ converges weakly to a probability measure R and $(Z_i)_{i\in\mathbb{N}}$ are $\mathbb{R}^d$-valued independent and identically distributed random variables having some exponential moments, *i.e.*

$$Ee^{\alpha|Z|}<+\infty \quad \text{for some} \quad 0<\alpha<+\infty.$$

The main improvement of this work is the relaxation of the steepness assumption concerning the cumulant generating function of the variables $(Z_i)_{i\in\mathbb{N}}$. In fact, Gärtner-Ellis' theorem is no longer available in this situation. As an application, we derive a LDP for the family of empirical measures $\frac{1}{n}\sum_1^n Z_i\delta_{x_i^n}$.
These measures are of interest in estimation theory (see Gamboa *et al.* [7], [12], Csiszar *et al.*, [6]), gas theory (see Ellis *et al.* [11], van den Berg *et al.* [22]), etc. We also derive LDPs for empirical processes in the spirit of Mogul'skii's theorem. Various examples illustrate the scope of our results.

Keywords Large deviations, empirical means, empirical measures, maximum entropy on the mean.

AMS subject classification (2000) 60F10, 60G57

Submitted to EJP on July 23, 2001. Final version accepted on October 12, 2001.

1 Introduction

Let $(Z_i)_{i\in\mathbb{N}}$ be a sequence of $\mathbb{R}^d$-valued independent and identically distributed (iid) random variables satisfying:

$$Ee^{\alpha|Z|} < +\infty \quad \text{for some} \quad \alpha > 0. \tag{1.1}$$

Let $(x_i^n,\ 1 \leq i \leq n,\ n \geq 1)$ be a $\mathcal{X}$-valued sequence of elements satisfying:

$$\frac{1}{n}\sum_1^n \delta_{x_i^n} \xrightarrow[n\to\infty]{weakly} R.$$

Here, R is assumed to be a strictly positive probability measure, that is $R(U) > 0$ whenever U is a nonempty open subset of $\mathcal{X}$. We prove in this article a Large Deviation Principle (LDP) for the weighted empirical mean:

$$\langle L_n, \mathbf{f}\rangle = \frac{1}{n}\sum_1^n \begin{pmatrix} f_1(x_i^n)\cdot Z_i \\ \vdots \\ f_m(x_i^n)\cdot Z_i \end{pmatrix} \quad \left(\stackrel{\triangle}{=} \frac{1}{n}\sum_1^n \mathbf{f}(x_i^n)\cdot Z_i\right),$$

where each f_k is a bounded continuous function from $\mathcal{X}$ to $\mathbb{R}^d$ and $\cdot$ denotes the scalar product in $\mathbb{R}^d$.
The main improvement of the paper is to remove the steepness assumption on the cumulant generating function of Z_i (for a definition of steepness, see [9], def. 2.3.5). Without this assumption, the Gärtner-Ellis theorem is no longer available. More precisely, one cannot expect to use Cramér's exponential change of measure technique to derive the Large Deviation lower bound. Our approach consists of using an exponential approximation to (loosely speaking) deduce the stated LDP from the sharp form of Cramér's theorem (see Bahadur and Zabell [2] and Section 6.1 in [9]) which holds true under condition (1.1). Similar LDPs are studied by Ben Arous, Dembo and Guionnet [1] in a context where the empirical measure $\frac{1}{n}\sum_1^n \delta_{x_i^n}$ is random. For related work concerning quadratic forms of gaussian processes, see Bercu, Gamboa and Lavielle [3], Bercu, Gamboa and Rouault [4], Gamboa, Rouault and Zani [13], Bryc and Dembo [5], Zani [23] and the references therein. In the case where the strict positivity of the probability measure R is not satisfied, counterexamples are built in Section 2.3 to show that the LDP can fail to happen.
As an application of the previous LDP, we derive various LDPs in infinite dimensional settings. We first establish the LDP for the following sequence of empirical measures:

$$L_n = \frac{1}{n}\sum_1^n Z_i\delta_{x_i^n}.$$

The rate function driving the LDP has the following form:

$$I(\mu) = \int \Lambda^*\left(\frac{d\mu_a}{dR}\right) dR + I^s(\mu_s),$$

where μ has the Lebesgue decomposition $\mu = \mu_a + \mu_s$ with $\mu_a \ll R$. This family is of interest in applications. It has been studied by Dacunha-Castelle and Gamboa in [7] and the LDP has been established by Gamboa and Gassiat in [12] to prove convergence results via the Maximum Entropy on the Mean (MEM) method in estimation theory. Following them, we shall call L_n the MEM empirical measure. Van den Berg, Dorlas, Lewis and Pulé in [22] and Ellis, Gough and Pulé in [11] have also studied such LDPs in more physical settings. We improve the previous works in two directions. We consider $\mathbb{R}^d$-valued variables (Z_i) and we remove the steepness assumption which stands in both papers. We also show that the assumption of strict positivity of R is necessary. Finally, we describe a side-effect in Section 3.2 in the case where the support of R (the space $\mathcal{X}$) is not compact. The LDP for $(L_n)_{n\geq 1}$ is also present in the book of Dembo and Zeitouni ([9], Section 7.2) where it is used to derive other results among which Mogul'skii's theorem. It is established under the following strong moment assumption:

$$Ee^{\alpha|Z|} < +\infty \quad \text{for all} \quad \alpha > 0. \tag{1.2}$$

It is interesting to note that no additional term involving μ_s appears under condition (1.2), *i.e.* $I(\mu) = \int \Lambda^*(d\mu/dR)\, dR$ if $\mu \ll R$ and ∞ otherwise. In the context of Sanov's theorem, the relation between exponential moment assumptions and the appearance of a singular term has been studied by Léonard and the author in [15].
The MEM Large Deviation Principle is then used to derive Mogul'skii type theorems. Namely LDPs are derived for the random functions:

$$t \mapsto \bar{Z}_n(t) = \frac{1}{n}\sum_{i=1}^{[nt]} Z_i \quad \text{and} \quad t \mapsto \tilde{Z}_n(t) = \bar{Z}_n(t) + \left(t - \frac{[nt]}{n}\right) Z_{[nt]+1}.$$

Following Lynch and Sethuraman [16] (see also de Acosta [8]), the LDPs are derived in bv, the set of bounded variation functions endowed with the weak-$*$ topology $\sigma(bv, C([0,1],\mathbb{R}^d))$. Under condition (1.2), the LDPs are well-known (see [9] for an updated account). Under condition (1.1), the LDP has been established by Mogul'skii [17] and the interesting form of the rate function is due to Lynch and Sethuraman [16]. For related work concerning processes with independent increments and satisfying condition (1.1), see de Acosta [8], Léonard [14], Mogul'skii [18].
The paper is organized as follows. The LDP for the weighted empirical mean is stated in Section 2. The LDP for the MEM empirical measure is established in Section 3. Mogul'skii type theorems are derived in Section 4. In Section 5 we prove the LDP for the weighted empirical mean. Finally, Sections 6 and 7 are devoted to additional proofs related to examples.

2 The LDP for the weighted empirical mean

We introduce here some notations and the main assumptions that hold all over the paper.

2.1 Notations and Assumptions

Let $\mathcal{X}$ be a topological vector space endowed with its Borel σ-field $\mathcal{B}(\mathcal{X})$ and let R be a probability measure on $\mathcal{X}$. We denote by $C(\mathcal{X})$ (resp. $C_d(\mathcal{X})$) the set of $\mathbb{R}$-valued (resp. $\mathbb{R}^d$-valued) continuous bounded functions on $\mathcal{X}$, by $L^1_d(\mathcal{X})$ the set of $\mathbb{R}^d$-valued R-integrable functions on $\mathcal{X}$ and by $\mathcal{P}(\mathcal{X})$ the set of probability measures on $\mathcal{X}$. We shall sometimes drop $\mathcal{X}$ and denote the previous sets by C, C_d, L^1_d.

Let $|\cdot|$ denote a norm on any finite-dimensional vector space (usually $\mathbb{R}$, $\mathbb{R}^d$ or $\mathbb{R}^{m\times d}$). We denote by $\|\cdot\|$ the supremum norm on the space of bounded continuous functions from $\mathcal{X}$ with values in $\mathbb{R}$, $\mathbb{R}^d$ or $\mathbb{R}^{m\times d}$, *i.e.* $\|f\| = \sup_{x\in\mathcal{X}} |f(x)|$. As usual, δ_a is the Dirac measure at a. We shall make the following assumptions:

Assumption A-1 *The family* $(x_i^n)_{1\le i\le n, n\ge 1} \subset \mathcal{X}$ *satisfies*

$$\frac{1}{n}\sum_1^n \delta_{x_i^n} \xrightarrow[n\to\infty]{\text{weakly}} R \quad \textit{where} \quad R \in \mathcal{P}(\mathcal{X}). \tag{2.1}$$

Assumption A-2 *R is a probability measure on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ satisfying*

if U is a non-empty open set then $R(U) > 0$.

Remark 2.1 The combination of Assumptions (A-1) and (A-2) is standard (see [1], [11] and [12]). In Section 2.3, we derive counterexamples in the case where Assumption (A-2) is not fulfilled. □

Assumption A-3 *Let $(Z_i)_{i\in\mathbb{N}}$ be a sequence of $\mathbb{R}^d$-valued independent and identically distributed random variables with distribution $\mathbb{P}_Z$. The following exponential moment condition holds:*

$$Ee^{\alpha|Z|} < +\infty \quad \textit{for some} \quad \alpha > 0. \tag{2.2}$$

Here, $\mathbb{P}_Z$ is the image of $\mathbb{P}$, that is $\mathbb{P}\{Z_i \in A\} = \mathbb{P}_Z(A)$, where $Z_i : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Furthermore, we denote by Λ the cumulant generating function of Z and by Λ^* its convex conjugate:

$$\begin{aligned} \Lambda(\lambda) &= \log Ee^{\lambda\cdot Z} \quad \text{for } \lambda \in \mathbb{R}^d, \\ \Lambda^*(z) &= \sup_{\lambda\in\mathbb{R}^d} \{\lambda\cdot z - \Lambda(\lambda)\} \quad \text{for } z \in \mathbb{R}^d, \end{aligned}$$

where $\cdot$ denotes the scalar product in $\mathbb{R}^d$. As usual, $\mathcal{D}_\Lambda = \{\lambda \in \mathbb{R}^d,\ \Lambda(\lambda) < \infty\}$ is the effective domain of Λ.

Let $\mathbf{a}$ be a $m\times d$ matrix and let $z \in \mathbb{R}^d$. We denote by $\cdot$ the usual matrix product, that is

$$\mathbf{a}\cdot z = \begin{pmatrix} a_1\cdot z \\ \vdots \\ a_m\cdot z \end{pmatrix},$$

where a_j is the j^{th} row of the matrix $\mathbf{a}$. Hence, $\cdot$ denotes the scalar product $\lambda\cdot z$ or the matrix product $\mathbf{a}\cdot z$, depending on the context. Let $\mathbf{f} : \mathcal{X} \to \mathbb{R}^{m\times d}$ be a (matrix-valued) bounded continuous function, then

$$\mathbf{f}(x)\cdot z = \begin{pmatrix} f_1(x)\cdot z \\ \vdots \\ f_m(x)\cdot z \end{pmatrix},$$

where each $f_j \in C_d(\mathcal{X})$ is the j^{th} row of the matrix $\mathbf{f}$. Let $u : \mathcal{X} \to \mathbb{R}^d$ be a measurable function, we denote by

$$\int_{\mathcal{X}} \mathbf{f}(x)\cdot u(x)\, R(dx) = \begin{pmatrix} \int_{\mathcal{X}} f_1(x)\cdot u(x)\, R(dx) \\ \vdots \\ \int_{\mathcal{X}} f_m(x)\cdot u(x)\, R(dx) \end{pmatrix}.$$

We now introduce the weighted empirical mean and study it in the sequel.

$$\langle L_n, \mathbf{f}\rangle \stackrel{\triangle}{=} \frac{1}{n}\sum_{i=1}^{n} \mathbf{f}(x_i^n)\cdot Z_i.$$

We shall follow the convention that $x \in \mathcal{X}$, y and θ are elements of $\mathbb{R}^m$ and z and λ, of $\mathbb{R}^d$.

2.2 Statement of the LDP

We state here the main result of the article. Both results of Section 3 (LDP for empirical measures) and results of Section 4 (LDPs for random functions) rely on it.

Theorem 2.2 *Let $\mathbf{f} : \mathcal{X} \to \mathbb{R}^{m\times d}$ be a continuous bounded function. Assume that (A-1), (A-2) and (A-3) hold. Then the family*

$$\langle L_n, \mathbf{f}\rangle = \frac{1}{n}\sum_{1}^{n} \mathbf{f}(x_i^n)\cdot Z_i$$

satisfies the large deviation principle in $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ with the good rate function

$$I_{\mathbf{f}}(y) = \sup_{\theta\in\mathbb{R}^m}\left\{\theta\cdot y - \int_{\mathcal{X}} \Lambda\Big[\sum_{j=1}^{m}\theta_j f_j(x)\Big]\, R(dx)\right\} \quad \text{for } y \in \mathbb{R}^m, \tag{2.3}$$

where $f_j \in C_d(\mathcal{X})$ denotes the j^{th} row of the matrix $\mathbf{f}$.

The proof of Theorem 2.2 is postponed to Section 5.

Remark 2.3 Note that the rate function $I_{\mathbf{f}}$ is expressed as the convex conjugate of $\int_{\mathcal{X}} \Lambda[\sum_1^m \theta_j f_j(x)]\, R(dx)$ which plays the rôle of the limiting cumulant generating function. □

2.3 Counterexamples when Assumption A-2 is not fulfilled

Consider the following empirical probability measures over the space $[0,2]$:

$$R_n = \frac{1}{n}\sum_{i=1}^{n-1}\delta_{\frac{i}{n}} + \frac{\delta_2}{n}.$$

Then R_n (weakly) converges to $\ell(dx)$ where ℓ denotes the Lebesgue measure on $[0,1]$. As a measure on $[0,2]$, ℓ does not satisfy (A-2). We build in this section a probability measure $\mathbb{P}_Z$ and a sequence of iid $\mathbb{P}_Z$-distributed random variables Z_i satisfying Assumption (A-3) such that:

• if f is a continuous function whose values are 0 on $[0,1]$ and 1 on 2, then the random variables

$$\frac{1}{n}\sum_{i=1}^{n-1} f\left(\frac{i}{n}\right) Z_i + \frac{f(2)}{n} Z_n = \frac{Z_n}{n}$$

do not satisfy a LDP.

• if f is a continuous function whose values are 1 on $[0,1]$ and -1 on 2, then the random variables

$$\frac{1}{n}\sum_{i=1}^{n-1} f\left(\frac{i}{n}\right) Z_i + \frac{f(2)}{n} Z_n = \frac{1}{n}\sum_{i=1}^{n-1} Z_i - \frac{Z_n}{n}$$

do not satisfy a LDP.

The first counterexample shows that a particle can fail to satisfy a LDP and illustrates the regularizing effect of the mean: Though $\frac{Z}{n}$ does not satisfy a LDP, $\frac{1}{n}\sum_1^n Z_i$ satisfies a LDP (Cramér's theorem). We shall prove in Lemma 5.1 that $\frac{1}{n}\sum_1^{N(n)} Z_i$ holds as soon as $N(n)/n \to \rho > 0$.

The second counterexample, whose study is much more involved, shows that even if $\frac{1}{n}\sum_1^{n-1} Z_i$ satisfies a LDP, the addition of a single contribution $\frac{Z}{n}$ can break the LDP and illustrates the fact that "without all the exponential moments, a single particle can modify a LDP".

These remarks give us a better understanding of Assumption (A-2) (strict positivity of R). In fact, assume that $\frac{Z}{n}$ does not satisfy a LDP and consider $\frac{1}{n}f(x_i)Z_i$. Assumption (A-2) ensures that there cannot be, around any point $x_0 \in \mathcal{X}$, a single particle which would break the LDP. In fact, there are enough points around x_0 to ensure the regularizing effect of the mean. Let us explain the phenomenon. Let V be a neighborhood of x_0 on which f is almost constant (recall that f is continuous): $f \approx a$. Then,

$$\begin{aligned}\frac{1}{n}\sum_1^n f(x_i)Z_i &= \frac{1}{n}\sum_{x_i\in V} f(x_i)Z_i + \frac{1}{n}\sum_{x_i\notin V} f(x_i)Z_i \\ &\approx \frac{a}{n}\sum_{x_i\in V} Z_i + \cdots .\end{aligned}$$

Let us compute the number of x_i's which belong to V:

$$N_V(n) = \sharp\{x_i \in V\} = \sum_1^n 1_V(x_i) \quad \text{and} \quad \frac{N_V(n)}{n} \to R(V) > 0,$$

where the last limit follows from (A-2) as soon as $R(\partial V) = 0$. This ensures that around x_0, $\frac{1}{n}\sum_{x_i \in V} Z_i$ satisfies a LDP (recall Lemma 5.1, previously mentionned). Thus, the phenomenon encountered with the second counterexample cannot hold. We shall first study the case of the "single particle" $\frac{Z}{n}$.

2.3.1 Example of a particle which does not satisfy a LDP

Consider the random variable Z with distribution

$$\mathbb{P}_Z(a_n) = \mathbb{P}\{Z = a_n\} = C\frac{e^{-a_n}}{1 + a_n^3} \quad \text{for } n \geq 0,$$

where $a_n = 16^n$ and C is a normalizing constant. This probability measure has two interesting features: It is "lacunary" and the rate function $\Lambda^*(z)$ is linear for z large enough (see Proposition 6.1 and its proof). We first show that:

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left\{\frac{Z}{n} \in]1, 2[\right\} = -\infty. \tag{2.4}$$

In fact, consider the subsequence $\phi(n) = 2a_n$ then

$$\mathbb{P}\{Z/\phi(n) \in]1, 2[\} = \mathbb{P}\{\phi(n) < Z < 2\phi(n)\} = 0,$$

as $\phi(n) = 2a_n > a_n$ and $2\phi(n) = 4a_n < a_{n+1} = 16a_n$. Therefore, (2.4) is proved. We show now that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left\{\frac{Z}{n} \in [1 + \eta, 2 - \eta]\right\} \geq -\frac{8}{7}. \tag{2.5}$$

Consider the subsequence $\phi(n) = 14a_{n-1}$ then

$$\mathbb{P}\left\{\frac{Z}{\phi(n)} \in [1 + \eta, 2 - \eta]\right\} = \mathbb{P}\{Z = a_n\} = c\frac{e^{-a_n}}{1 + a_n^3}.$$

Therefore $\lim_n 1/\phi(n) \log \mathbb{P}\{Z/\phi(n) \in [1 + \eta, 2 - \eta]\} = -8/7$ and (2.5) is proved. Assume now that $\frac{Z}{n}$ satisfies a LDP with rate function J, then:

$$-8/7 \leq \limsup_n \frac{1}{n} \log \mathbb{P}\{Z/n \in [1 + \eta, 2 - \eta]\} \leq -\inf_{u\in[1+\eta,2-\eta]} J(u).$$

Moreover,

$$-\inf_{u\in]1,2[} J(u) \leq \liminf_n \frac{1}{n} \log \mathbb{P}\{Z/n \in]1, 2[\} \leq -\infty.$$

Necessarily, one should have $\inf_{u\in]1,2[} J(u) = +\infty > 8/7 \geq \inf_{u\in[1+\eta,2-\eta]} J(u)$, which is impossible. Hence $\frac{Z}{n}$ does not satisfy a LDP.

2.3.2 The second counterexample

Let Z_i be iid and $\mathbb{P}_Z$-distributed (as defined previously). We show here that $1/n\sum_1^{n-1} Z_i - Z_n/n$ does not satisfy a LDP. The counterexample is based on the following idea: Consider the event $\{\hat{Z}_n - Z_n/n \in [z-2\,\epsilon, z+2\,\epsilon]\}$ where z is negative and where $\hat{Z}_n = 1/n\sum_1^{n-1} Z_i$. As $\hat{Z}_n$ is always positive, the only way the previous event is realized is by Z_n/n being large. As the "particle" Z_n/n does not satisfy a LDP, the same should hold for $\hat{Z}_n - Z_n/n$.

Proposition 2.4 *The random variables* $(1/n\sum_{i=1}^{n-1} Z_i - Z_n/n)$ *do not satisfy a LDP.*

The proof of this proposition, technically involved, is postponed to Section 6.

3 The LDP for MEM empirical measures

We establish here the LDP for empirical measures.

3.1 The LDP

In Theorem 3.1, we state a LDP for empirical measures. We first introduce a few notations. Denote by $C_d^*(\mathcal{X})$ the topological dual of $C_d(\mathcal{X})$. We endow it with the weak-$*$ topology $\mathcal{E} = \sigma(C_d^*, C_d)$ and with the associated Borel σ-field $\mathcal{B}(C_d^*)$. If $\xi \in C_d^*(\mathcal{X})$ and $f \in C_d(\mathcal{X})$, we shall denote by $\langle \xi, f\rangle$ the duality bracket.

Assumption A-4 *$\mathcal{X}$ is a compact Hausdorff space.*

In the case where $\mathcal{X}$ is compact, the continuous linear forms over $C_d(\mathcal{X})$ are vector measures. We denote by $M^d(\mathcal{X})$ the set of vector measures with value in $\mathbb{R}^d$, that is $\mu \in M^d(\mathcal{X})$ iff $\mu = (\mu_1, \dots, \mu_d)$ where each $\mu_i \in M(\mathcal{X})$. Let $f \in C_d(\mathcal{X})$ and $\mu \in M^d(\mathcal{X})$. We denote by

$$\int_{\mathcal{X}} f(x) \cdot \mu(\,dx) \triangleq \sum_{l=1}^{d} \int_{\mathcal{X}} f_l(x)\mu_l(\,dx). \tag{3.1}$$

Let $\mathbf{f} : \mathcal{X} \to \mathbb{R}^{m\times d}$ be a (matrix valued) bounded continuous function and let $f_i \in C_d(\mathcal{X})$ be the j^{th} row of the matrix $\mathbf{f}$. We denote by

$$\int \mathbf{f} \cdot d\mu \triangleq \begin{pmatrix} \int_{\mathcal{X}} f_1(x) \cdot \mu(dx) \\ \vdots \\ \int_{\mathcal{X}} f_m(x) \cdot \mu(dx) \end{pmatrix}. \tag{3.2}$$

We shall endow $M^d(\mathcal{X})$ with the weak-$*$ topology $\Sigma = \sigma(M^d, C_d)$ which makes every linear form $\Gamma_f : \mu \mapsto \int f \cdot d\mu$ continuous and with the associated Borel σ-field $\mathcal{B}(M^d)$. Let $\mu \in M^d(\mathcal{X})$. We denote by μ_a its absolutely continuous part with respect to R and by μ_s its singular part.

Theorem 3.1 *Assume that (A-1), (A-2) and (A-3) hold. Then the family*

$$L_n = \frac{1}{n}\sum_1^n Z_i \delta_{x_i^n}$$

satisfies a large deviation principle in $(C_d^*(\mathcal{X}), \mathcal{E}, \mathcal{B}(C_d^*))$ *with the good rate function*

$$I(\xi) = \sup_{f\in C_d(\mathcal{X})} \{\langle \xi, f\rangle - \int_{\mathcal{X}} \Lambda[f(x)]\, R(dx)\}.$$

Assume moreover that (A-4) holds. Then the LDP holds in $(M^d(\mathcal{X}), \Sigma, \mathcal{B}(M^d))$ *with the good rate function*

$$\begin{aligned} I(\mu) &= \sup_{f\in C_d(\mathcal{X})} \{\int_{\mathcal{X}} f(x)\cdot \mu(dx) - \int_{\mathcal{X}} \Lambda[f(x)]\, R(dx)\} \\ &= \int_{\mathcal{X}} \Lambda^*\Big[\frac{d\mu_a}{dR}(x)\Big]\, R(dx) + \int_{\mathcal{X}} \rho\Big[\frac{d\mu_s}{d\theta}(x)\Big]\, \theta(dx), \end{aligned} \tag{3.3}$$

where $\rho(z) = \sup\{\lambda\cdot z,\ \lambda\in \mathcal{D}_\Lambda\}$ *is the recession function of* Λ^* *and* θ *is any real-valued nonnegative measure with respect to which* μ_s *is absolutely continuous.*

Remark 3.2 [Rôle of θ] If the Z_i's are $\mathbb{R}$-valued, one can choose $\theta = |\mu_s| = \mu_s^+ + \mu_s^-$ (where $\mu_s = \mu_s^+ - \mu_s^-$ by the Hahn-Jordan decomposition) and the rate function is given by:

$$I(\mu) = \int \Lambda^*\Big[\frac{d\mu_a}{dR}\Big]\, dR + \rho(1)\mu_s^+(\mathcal{X}) + \rho(-1)\mu_s^-(\mathcal{X}).$$

In the $\mathbb{R}^d$-case, one can choose $\theta = \sum_1^d |\mu_s^k|$ where $\mu_s = (\mu_s^1, \dots, \mu_s^d)$. □

Remark 3.3 [Compactness of $\mathcal{X}$] Without the compactness assumption, the continuous linear functionals on $C_d(\mathcal{X})$ are no longer measures. They are regular bounded additive set functions [10]. Moreover, Assumption (A-4) is central to identify the rate function (see (3.3)) via an identity due to Rockafellar [21]. □

Remark 3.4 In the case where $Ee^{\alpha|Z|} < \infty$ for every $\alpha > 0$, the recession function is $\rho(z) = \sup\{\lambda\cdot z,\ \lambda\in\mathbb{R}^d\} = +\infty$ everywhere except in zero where $\rho(0) = 0$. Hence $I(\mu)$ is finite only if μ is absolutely continuous with respect to R and the rate function is:

$$I(\mu) = \int_{\mathcal{X}} \Lambda^*\Big[\frac{d\mu}{dR}(x)\Big]\, R(dx) \text{ if } \mu \ll R,\ +\infty \text{ otherwise.}$$

This is in accordance with Theorem 7.2.3 in [9] where no extra term involving μ_s appears.
□

Proof • *The LDP.* Denote by $C'_d(\mathcal{X})$ (resp. $C'(\mathcal{X})$) the algebraic dual of $C_d(\mathcal{X})$ (resp. $C(\mathcal{X})$). Denote by $\langle\ ,\ \rangle$ the duality bracket between these spaces and consider the mapping

$$p_{f_1,\dots,f_m} : C'_d(\mathcal{X}) \to \mathbb{R}^m, \quad \xi \mapsto \begin{pmatrix} \langle \xi, f_1 \rangle \\ \vdots \\ \langle \xi, f_m \rangle \end{pmatrix},$$

where $f_i \in C_d(\mathcal{X})$. Then $p_{f_1,\dots,f_m}(L_n) = \langle L_n, \mathbf{f} \rangle$ satisfies a LDP by Theorem 2.2. By Dawson-Gärtner's theorem, L_n satisfies a LDP in $C'_d(\mathcal{X})$ endowed with the weak-$*$ topology with the good rate function

$$\begin{aligned} I(\xi) &= \sup_{m\geq 1} \sup_{f_1,\dots,f_m \in C_d(\mathcal{X})} \sup_{\theta\in\mathbb{R}^m} \{\sum_1^m \theta_i \langle \xi, f_i \rangle - \int_{\mathcal{X}} \Lambda[\sum_1^m \theta_i f_i(x)]\, R(dx)\} \\ &= \sup_{f\in C_d(\mathcal{X})} \{\langle \xi, f\rangle - \int_{\mathcal{X}} \Lambda[f(x)]\, R(dx)\} \quad \text{for} \quad \xi \in C'_d(\mathcal{X}). \end{aligned}$$

• *Restriction of the LDP.* Let us show that $I(\xi) < +\infty$ implies that ξ is a continuous linear form (*i.e.* $\xi \in C^*_d(\mathcal{X})$). Assume that $I(\xi) < +\infty$. Then for all $f \in C_d(\mathcal{X})$, $f \neq 0$

$$\langle \xi, \frac{f}{a\|f\|} \rangle \leq I(\xi) + \int \Lambda\Big(\frac{f}{a\|f\|}\Big)\, dR \leq I(\xi) + \Lambda_{|Z|}\Big(\frac{1}{a}\Big), \tag{3.4}$$

where $\Lambda_{|Z|}$ denotes the cumulant generating function of $|Z|$. For a large enough $\Lambda_{|Z|}\left(\frac{1}{a}\right)$ is finite by (A-3) and $\langle \xi, f \rangle \leq K\|f\|$. Considering $-f$, we get $|\langle \xi, f\rangle| \leq K\|f\|$. Thus ξ is a continuous linear form and the LDP holds in the stated space by Lemma 4.1.5 in [9]. If moreover Assumption (A-4) holds, Riesz's representation theorem implies that ξ can be represented as a $\mathbb{R}^d$-valued measure over $\mathcal{X}$, *i.e.* $\xi \in M^d(\mathcal{X})$. We shall denote it by μ. We can now apply Lemma 4.1.5 in [9] to obtain the LDP in $(M^d(\mathcal{X}), \Sigma, \mathcal{B}(M^d))$.
• *Representation of the rate function.* Under (A-2) and (A-4), Theorem 5 in [21] yields

$$I(\mu) = \int_{\mathcal{X}} \Lambda^*\Big[\frac{d\mu_a}{dR}(x)\Big]\, R(dx) + \int_{\mathcal{X}} \rho\Big[\frac{d\mu_s}{d\theta}\Big]\, d\theta,$$

where ρ is the recession function of Λ^* and θ is any real-valued nonnegative measure with respect to which μ_s is absolutely continuous. As Λ is the convex conjugate of Λ^*, ρ is the support function of Λ ([19], theorem 13.3), that is:

$$\rho(z) = \sup\{\lambda \cdot z,\ \lambda \in \mathcal{D}_\Lambda\}.$$

Hence Theorem 3.1 is proved. □

3.2 Side-effects in the noncompact case: A fact and a heuristic

Assume that R is a probability measure on $\mathbb{R}^+$ satisfying (A-2) and that there exists a sequence (x_i^n) satisfying (A-1). Let (Z_i) be iid $\mathbb{P}_Z$-distributed $\mathbb{R}$-valued nonnegative random variables where $\mathbb{P}_Z(\,dz) = 1_{[0,\infty)}(z)e^{-z}\,dz$. In this case (A-3) is satisfied and L_n satisfies a LDP in $(C^*_d(\mathbb{R}^+), \mathcal{E}, \mathcal{B}(C^*_d))$ by Theorem 3.1.

The fact

Consider $\mu_n = a\,R + \delta_n$ then $I(\mu_n) \le 1 + \Lambda^*(a)$. In fact,

$$\begin{aligned} I(\mu_n) &= \sup_{f \in C(\mathbb{R}^+)} \left\{ \int a\,f\,dR + f(n) - \int \Lambda(f)\,dR \right\} \\ &= \sup_{f \in C(\mathbb{R}^+),\ f \le 1} \left\{ \int a\,f\,dR + f(n) - \int \Lambda(f)\,dR \right\}, \end{aligned}$$

where the last equality follows from the fact that $\int \Lambda(f)\,dR = +\infty$ if $f > 1$ (one can check that $\Lambda(\lambda) = +\infty$ if $\lambda > 1$). Finally if $f \le 1$ then:

$$\begin{aligned} \int a\,f\,dR + f(n) - \int \Lambda(f)\,dR &\le \int a\,f\,dR - \Lambda\left(\int f\,dR\right) + 1 \\ &\le \Lambda^*(a) + 1. \end{aligned}$$

Thus $I(\mu_n) \le \Lambda^*(a) + 1$. Therefore $\mu_n \in \{I \le \Lambda^*(a) + 1\}$ which is a compact set. Hence μ_n admits cluster points (as limits of converging subnets since $C_d^*(\mathbb{R}^+)$ is not metrizable) which obviously are not measures.

The heuristic

A heuristical interpretation of the previous remark is the following: In a large deviation regime, L_n can behave asymptotically as $a\,R + \delta_n$. This legitimates the appearance of linear forms which are no longer measures in the non-compact case. To see this, first note that the "particle" $\frac{Z}{n}$ satisfies a LDP with good rate function

$$\delta^*(z) = \begin{cases} z & \text{if } z \ge 0 \\ +\infty & \text{else} \end{cases}.$$

Therefore the contribution of one particle is of importance in the LD phenomenon. Since $\lim_n 1/n \sum \delta_{x_i^n} = R$ which is spread over $\mathbb{R}^+$, there exists a subsequence of (x_i^n), say x_n, satisfying $\lim_n x_n = +\infty$. Consider now

$$L_n = \frac{1}{n} \sum_{i=1}^{n-1} Z_i\ \delta_{x_i^n} + \frac{Z_n}{n} \delta_{x_n}.$$

A large deviation behaviour can occur with $\frac{Z_n}{n}$ being large, the rest of the empirical measure being standard. This heuristic yields $L_n \approx m\,R + \delta_{x_n}$ where $m = EZ$.

3.3 Example: Lack of strict convexity for the rate function I

In this example, we shall consider the following setup: $\mathcal{X} = [0,1]$, $R(\,dx) = \ell(dx)$ where $\ell(dx)$ stands for the Lebesgue measure on $[0,1]$, $(x_i^n) = (i/n)$ and $\mathbb{P}_Z(\,dz) = C.1_{[0,\infty)}(z) \frac{e^{-z}}{1+z^3}\,dz$. Thus (Z_i) are nonnegative real valued random variables. Let

$z^* = \Lambda'(1)$ and $m = EZ$. Standard considerations yield that Λ is not steep and Λ^* is linear for $z > z^*$:

$$\Lambda^*(z) = z - z^* + \Lambda^*(z^*) \quad \text{for} \quad z > z^*. \tag{3.5}$$

In this context, the recession function is $\rho(z) = z$ and it can easily be shown that the rate function $I(\mu)$ of the MEM large deviation principle is finite only if μ is a positive measure. Therefore, I has the form:

$$I(\mu) = \int_{[0,1]} \Lambda^* \left[\frac{d\mu_a}{d\ell}(x) \right] \ell(dx) + \mu_s[0,1].$$

if μ is a positive measure and $I(\mu) = +\infty$ otherwise.

3.3.1 The value of the rate function I for a special measure

Let κ be the Cantor function on $[0,1]$, that is κ is continuous, non-decreasing from $[0,1]$ to $[0,1]$ and κ's derivative is ℓ-a.e. null. In particular, κ is the repartition function of a probability measure μ_κ singular with respect to ℓ and

$$I(\ell + \mu_\kappa) = \Lambda^*(1) + \mu_\kappa[0,1] = \Lambda^*(1) + 1.$$

3.3.2 Existence of several minimizers under a convex constraint

Consider the convex constraint $\mathcal{C} = \{\mu \in M([0,1]),\ \langle \mu, 1 \rangle = E\}$ and denote by $\mathcal{M}$ the set of minimizers of the rate function I under $\mathcal{C}$. Then the following holds:

Proposition 3.5 *Let E be a real number:*

1. *if $E \in [m, z^*)$, then there exists a unique minimizer of I under the constraint $\mathcal{C}$. This minimizer μ is defined by $\frac{d\mu}{d\ell}(x) = E$ and $I(\mu) = \Lambda^*(E)$.*

2. *if $E \geq z^*$, then every positive measure satisfying $\mu = \mu_a + \mu_s$ where $g(x) = \frac{d\mu_a}{d\ell}(x)$ satisfies $g(x) \geq z^* \quad \ell - a.e.$ and $\langle \mu, 1 \rangle = E$ is a minimizer. Moreover, $I(\mu) = E - z^* + \Lambda^*(z^*)$.*

The proof of Proposition 3.5 is postponed to Section 7.

Remark 3.6 In view of part 2 of the previous proposition, all the following measures belong to $\mathcal{M}$ in the case where $E > z^*$:

$$\alpha\ \ell(dx) + \beta\ \delta_u(dx) + \gamma\ \mu_\kappa(dx),$$

where $\alpha + \beta + \gamma = E$, $\alpha \geq z^*$ and $u \in [0,1]$. In particular, there are infinitely many minimizers. □

Remark 3.7 As a by-product of the second part of the proposition, the rate function I fails to be stricly convex on its domain $\mathcal{D}_I = \{\mu,\ I(\mu) < \infty\}$. A similar fact has been noticed in the context of Sanov's theorem in [15]. □

4 Mogul'skii type results

In this section, we derive functional LDPs for the random functions

$$t \mapsto \bar{Z}_n(t) = \frac{1}{n}\sum_{i=1}^{[nt]} Z_i \quad \text{and} \quad t \mapsto \tilde{Z}_n(t) = \bar{Z}_n(t) + \left(t - \frac{[nt]}{n}\right) Z_{[nt]+1},$$

where $[x]$ denotes the integer part of x. These results are essentially corollaries of the MEM large deviations principle previously derived in the case where $\mathcal{X} = [0,1]$, $R(dx) = \ell(dx)$ is the Lebesgue measure on $[0,1]$ and $(x_i^n) = \left(\frac{i}{n}\right)$. One can check that, in this situation, Assumptions (A-1) and (A-2) hold true.
Following Lynch and Sethuraman [16] (see also de Acosta [8] and Zani [23], chapter 4), we introduce some notations. Let $bv([0,1],\mathbb{R}^d)$ (shortened in bv) be the space of functions of bounded variation on $[0,1]$. We identify bv with $M^d([0,1])$ in the usual manner: To $f \in bv$ there corresponds μ^f characterized by $\mu^f([0,t]) = f(t)$. Up to this identification, $C_d([0,1])$ is the topological dual of bv. We endow bv with the weak-$*$ topology $\sigma(bv, C_d([0,1]))$ (shortened in σ_w) and with the associated Borel σ-field $\mathcal{B}_w$. Let $f \in bv$ and μ^f be the associated measure in $M^d([0,1])$. Consider the Lebesgue decomposition of μ^f, $\mu^f = \mu_a^f + \mu_s^f$ where μ_a^f denotes the absolutely continuous part of μ^f with respect to dx and μ_s^f its singular part. We denote by $f_a(t) = \mu_a^f([0,t])$ and by $f_s(t) = \mu_s^f([0,t])$.

4.1 The LDP for the discontinuous line $\bar{Z}_n(.)$

The LDP for the discontinuous line can be found in [9] in the case where $\Lambda(\lambda) = \ln Ee^{\lambda\cdot Z_i} < \infty$ for all $\lambda \in \mathbb{R}^d$. It is established under the supremum norm topology. In the following theorem, we loosen the assumptions on the exponential moments of Z_i. However, the LDP is stated under a weaker topology.

Theorem 4.1 *Assume that (A-3) holds. Then the random functions $\left(\bar{Z}_n(t)\right)_{t\in[0,1]}$ satisfy the LDP in $(bv, \sigma_w, \mathcal{B}_w)$ with the good rate function*

$$\Phi(f) = \int_{[0,1]} \Lambda^*(f_a'(t))\,dt + \int_{[0,1]} \rho(f_s'(t))\,d\theta(t),$$

where θ is any real-valued nonnegative measure with respect to which μ_s^f is absolutely continuous and $f_s' = d\mu_s^f/d\theta$.

Remark 4.2 Note that the definition of f_s' is θ-dependent. See also Remark 3.2. $\square$

Proof consider $\Pi : M^d \to bv$ where

$$\Pi(\mu) = \begin{pmatrix} \mu_1[0,t] \\ \vdots \\ \mu_d[0,t] \end{pmatrix}_{t\in[0,1]}.$$

and recall that $L_n = 1/n \sum_1^n Z_i \delta_{x_i}$. Then Π is a continuous bijection and $\Pi(L_n) = \bar{Z}_n$. As L_n satisfies the LDP with the good rate function (3.3) by Theorem 3.1, the contraction principle yields the LDP. $\square$

4.2 The LDP for the polygonal line $\tilde{Z}_n(.)$

The LDP for the polygonal line $\tilde{Z}_n(.)$ has been established by Mogul'skii in [17]. Our results differ from his. In fact, our LDP is derived under a weaker topology than in [17]. However, both the state space and the rate function are explicit here.

Theorem 4.3 *Assume that (A-3) holds. Then the random functions* $\left(\tilde{Z}_n(t)\right)_{t\in[0,1]}$ *satisfy the LDP in* $(bv, \sigma_w, \mathcal{B}_w)$ *with good the rate function*

$$\Phi(f) = \int_{[0,1]} \Lambda^*(f_a'(t))\,dt + \int_{[0,1]} \rho(f_s'(t))\,d\theta(t),$$

where θ is any real-valued nonnegative measure with respect to which μ_s^f is absolutely continuous and $f_s' = d\mu_s^f/d\theta$.

Proof Let $f \in C_d([0,1])$ and consider the empirical measure $\tilde{L}_n$ characterized by

$$\langle \tilde{L}_n, f\rangle = \sum_{i=1}^n \int_{(i-1)/n}^{i/n} f(t)\,dt \cdot Z_i,$$

where the integral $\int_{(i-1)/n}^{i/n} f(t)\,dt$ is $\mathbb{R}^d$-valued. Then $\Pi(\tilde{L}_n) = \tilde{Z}_n$ (where Π is defined as in the previous proof) and the theorem is proved as far as we prove the LDP for $\tilde{L}_n$ with good rate function given by (3.3). To this end, let $\mathbf{f} : \mathcal{X} \to \mathbb{R}^{m\times d}$ be a (matrix valued) bounded continuous function and let $f_j \in \mathbb{R}^d$ be the j^{th} row of $\mathbf{f}$. we introduce $\langle \tilde{L}_n, \mathbf{f}\rangle$ defined by

$$\langle \tilde{L}_n, \mathbf{f}\rangle = \sum_1^n \begin{pmatrix} \int_{(i-1)/n}^{i/n} f_1(t)\,dt \cdot Z_i \\ \vdots \\ \int_{(i-1)/n}^{i/n} f_m(t)\,dt \cdot Z_i \end{pmatrix} \quad \left(\stackrel{\triangle}{=} \sum_1^n \int_{(i-1)/n}^{i/n} \mathbf{f}(t)\,dt \cdot Z_i\right).$$

Let us show that $\langle L_n, \mathbf{f}\rangle$ and $\langle \tilde{L}_n, \mathbf{f}\rangle$ are exponentially equivalent:

$$\begin{aligned} |\langle \tilde{L}_n, \mathbf{f}\rangle - \langle L_n, \mathbf{f}\rangle| &\le \sum_{i=1}^n |Z_i| \left| \int_{(i-1)/n}^{i/n} \mathbf{f}(t)\,dt - \frac{\mathbf{f}(i/n)}{n} \right| \\ &\le \sum_{i=1}^n |Z_i| \int_{(i-1)/n}^{i/n} |\mathbf{f}(t) - \mathbf{f}(i/n)|\,dt. \end{aligned}$$

Let $\epsilon > 0$ be fixed. As $\mathbf{f}$ is uniformly continuous on $[0,1]$, $|\mathbf{f}(t) - \mathbf{f}(i/n)| \le \epsilon$ for $t \in [(i-1)/n, i/n]$ and for n large enough. Therefore,

$$|\langle \tilde{L}_n, \mathbf{f}\rangle - \langle L_n, \mathbf{f}\rangle| \le \frac{1}{n} \sum_1^n |Z_i|\ \epsilon$$

for n large enough and

$$\begin{aligned}&\limsup_n \frac{1}{n}\log \mathbb{P}\{|\langle \tilde{L}_n, \mathbf{f}\rangle - \langle L_n, \mathbf{f}\rangle| > \delta\}\\ \leq\ & \limsup_n \frac{1}{n}\log \mathbb{P}\{1/n \sum_1^n |Z_i|\ \epsilon > \delta\} \leq -\Lambda^*_{|Z|}\left(\frac{\delta}{\epsilon}\right).\end{aligned}$$

As by (A-3) $\lim_{\epsilon\to 0}\Lambda^*_{|Z|}\left(\frac{\delta}{\epsilon}\right) = +\infty$, $\langle \tilde{L}_n, \mathbf{f}\rangle$ and $\langle L_n, \mathbf{f}\rangle$ are exponentially equivalent. Thus $\langle \tilde{L}_n, \mathbf{f}\rangle$ satisfies the LDP in $\mathbb{R}^m$ with good rate function given by (2.3) and one can prove the LDP for $\tilde{L}_n$ as in the proof of Theorem 3.1. □

5 Proof of Theorem 2.2

There are two parts in the proof of Theorem 2.2. First we establish the LDP in Section 5.1, then we identify the rate function in Section 5.2.

5.1 The Large Deviation Principle

The proof relies on several preliminary results. Via a rescaled version of Cramér's theorem (Lemma 5.1), we establish the LDP for finite-range step functions (Lemma 5.2). Let $\mathbf{f}(x) = \sum_1^p \mathbf{a}_k 1_{A_k}(x)$, then

$$\langle L_n, \mathbf{f}\rangle \stackrel{\mathcal{D}}{=} \frac{1}{n}\sum_{i=1}^{N_{A_1}(n)} \mathbf{a}_1 \cdot Z_i^{(1)} + \cdots + \frac{1}{n}\sum_{i=1}^{N_{A_p}(n)} \mathbf{a}_p \cdot Z_i^{(p)}$$

satisfies the LDP principle. Finally, we show in Proposition 5.3 that $(\langle L_n, \mathbf{f}^p\rangle)_{p\geq 1}$ is an exponential approximation of $\langle L_n, \mathbf{f}\rangle$ where $(\mathbf{f}^p)$ is a well-chosen sequence of finite-range step functions. This step is the key point of the proof and yields the LDP for $\langle L_n, \mathbf{f}\rangle$.

Lemma 5.1 *Let $(Z_i)_{i\geq 1}$ satisfy Assumption (A-3). Assume further that $(N_A(n))_{n\geq 1}$ is a sequence of integers satisfying:*

$$\lim_{n\to\infty} \frac{N_A(n)}{n} = R_A > 0,$$

then $\left(\frac{1}{N}\sum_1^{N_A(n)} Z_i\right)$ satisfies the LDP with good rate function $I(z) = R_A\ \Lambda^*(\frac{z}{R_A})$.

Proof We denote by $\Lambda^*(B) = \inf_{z\in B}\Lambda^*(z)$. First, notice that

$$\bar{Z}_n^A = \frac{1}{N_A(n)}\sum_{i=1}^{N_A(n)} Z_i$$

satisfies the LDP with good rate function Λ^* and with speed $N_A(n)$, that is

$$\begin{aligned} -\Lambda^*(\overset{\circ}{B}) &\leq \liminf_{n\to\infty} \frac{1}{N_A(n)} \log \mathbb{P}\left\{\bar{Z}_n^A \in B\right\} \quad \text{and} \\ -\Lambda^*(\bar{B}) &\geq \limsup_{n\to\infty} \frac{1}{N_A(n)} \log \mathbb{P}\left\{\bar{Z}_n^A \in B\right\}, \end{aligned}$$

where $\overset{\circ}{B}$ (resp. $\bar{B}$) denotes the interior (resp. the closure) of B. Indeed, this is a direct application of Cramér's theorem in $\mathbb{R}^d$ (see for example [9], section 6).
Consider on the other hand the degenerate random variables α_n with distribution given by $\mathbb{P}\{\alpha_n = \frac{N_A(n)}{n}\} = 1$ and which are independent of $(\bar{Z}_n^A)$. It is straightforward to check that (α_n) satisfies the LDP with speed $N_A(n)$ and with good rate function given by

$$\delta(t|R_A) = \begin{cases} 0 & \text{if } t = R_A \\ +\infty & \text{else.} \end{cases}$$

Therefore, the couple $(\alpha_n, \bar{Z}_n^A)$ satisfies the LDP with speed $N_A(n)$, and with good rate function

$$\delta \oplus \Lambda^*(t, z) = \delta(t|R_A) + \Lambda^*(z)$$

(Lynch and Sethuraman [16], lemma 2.8). Hence, the contraction principle yields the LDP for $(\alpha_n \bar{Z}_n^A)$ with speed $N_A(n)$ and with good rate function

$$I(z) = \inf\{\delta \oplus \Lambda^*(t, z'),\ tz' = z\} = \Lambda^*\left(\frac{z}{R_A}\right).$$

Thus $(\frac{1}{n}\sum_{i=1}^{N_A(n)} Z_i)$ satisfies the LDP with speed n and with good rate function

$$R_A \Lambda^*(z/R_A)$$

and Lemma 5.1 is proved. □

Lemma 5.2 *Assume that (A-1) and (A-3) hold and let $(A_k)_{1\leq k\leq p}$ be a family of measurable sets of $\mathcal{X}$ such that $R(A_k) > 0$. Assume further that each A_k is a continuity set for R (i.e. $R(\partial A_k) = 0$ where $\partial A_k = \bar{A}_k - \mathring{A}_k$). Consider the step function*

$$\mathbf{f}(x) = \sum_{k=1}^{p} \mathbf{a}_k 1_{A_k}(x),$$

where each $\mathbf{a}_k$ is a $m \times d$ matrix. Then

$$\begin{aligned} \langle L_n, \mathbf{f}\rangle &= \frac{1}{n}\sum_1^n \mathbf{f}(x_i^n)\cdot Z_i \\ &= \frac{1}{n}\sum_1^n (\mathbf{a}_1 \cdot Z_i)\ 1_{A_1}(x_i^n) + \cdots + \frac{1}{n}\sum_1^n (\mathbf{a}_p \cdot Z_i)\ 1_{A_p}(x_i^n) \end{aligned}$$

satisfies the LDP in $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ with the good rate function

$$
\begin{aligned}
I_{\mathbf{f}}(y) &= \inf\Big\{\sum_1^p R(A_k)\Lambda^*(u_k), \quad \sum_1^p \mathbf{a}_k \cdot u_k\, R(A_k) = y,\ u_k \in \mathbb{R}^d\Big\} \\
&= \inf_{u \in L_d^1} \Big\{\int_{\mathcal{X}} \Lambda^*(u(x))\, R(dx),\ \int_{\mathcal{X}} \mathbf{f}(x) \cdot u(x)\, R(dx) = y\Big\}
\end{aligned}
\tag{5.1}
$$

for $y \in \mathbb{R}^m$.

Proof • Let $\mathbf{f}(x) = \sum_1^p \mathbf{a}_k 1_{A_k}(x)$ and consider the variable Z_i. Its coefficient is $\mathbf{f}(x_i^n)$ and as x_i^n can change with n, so does $\mathbf{f}(x_i^n)$. Hence, $\langle L_n, \mathbf{f}\rangle$ has the following form:

$$
\langle L_n, \mathbf{f}\rangle = \frac{1}{n}\sum_{i \in I_1(n)} \mathbf{a}_1 \cdot Z_i + \cdots + \frac{1}{n}\sum_{i \in I_p(n)} \mathbf{a}_p \cdot Z_i,
$$

where $I_k(n) = \{i \le n,\ x_i^n \in A_k\}$. Let $N_{A_k}(n) = \sharp\{i \le n,\ x_i^n \in A_k\} = \sum_{i=1}^n 1_{A_k}(x_i^n)$. There exist p independent families $\Big(\tilde{Z}_i^{(k)}, i \ge 1\Big)_{1 \le k \le p}$ of iid random variables having the same distribution as Z and such that the following equality holds in distribution:

$$
\langle L_n, \mathbf{f}\rangle \overset{\mathcal{D}}{=} \frac{1}{n}\sum_{i=1}^{N_{A_1}(n)} \mathbf{a}_1 \cdot \tilde{Z}_i^{(1)} + \cdots + \frac{1}{n}\sum_{i=1}^{N_{A_p}(n)} \mathbf{a}_p \cdot \tilde{Z}_i^{(p)} \quad \big(\triangleq \langle \tilde{L}_n, \mathbf{f}\rangle\big).
\tag{5.2}
$$

• First note that $\lim_{n\to\infty} \frac{N_{A_k}(n)}{n} = \lim_{n\to\infty} \frac{\sum_{i=1}^n 1_{A_k}(x_i^n)}{n} = R(A_k)$ as A_k is a continuity set for R. Therefore we can apply Lemma 5.1 to each sequence $\frac{1}{n}\sum_{i=1}^{N_{A_k}(n)} \tilde{Z}_i^{(k)}$. Moreover, the sequence

$$
(\zeta_n)_{n \ge 1} = \Big(\frac{1}{n}\sum_{i=1}^{N_{A_1}(n)} \tilde{Z}_i^{(1)}, \ldots, \frac{1}{n}\sum_{i=1}^{N_{A_p}(n)} \tilde{Z}_i^{(p)}\Big)_{n \ge 1}
$$

satisfies the LDP with the good rate function

$$
I(z_1, \ldots, z_p) = \sum_{k=1}^p R(A_k)\ \Lambda^*\big(\frac{z_k}{R(A_k)}\big) \quad \text{for } z_k \in \mathbb{R}^d
$$

(Lynch and Sethuraman [16], lemma 2.8). The contraction principle yields the LDP for $\langle \tilde{L}_n, \mathbf{f}\rangle$ with the good rate function

$$
\begin{aligned}
I_{\mathbf{f}}(y) &= \inf\Big\{\sum_{k=1}^p R(A_k)\ \Lambda^*\Big(\frac{z_k}{R(A_k)}\Big),\ \sum_1^p \mathbf{a}_k \cdot z_k = y,\ z_k \in \mathbb{R}^d \quad \text{for } 1 \le k \le p\Big\} \\
&= \inf\Big\{\sum_1^p R(A_k)\ \Lambda^*(u_k),\ \sum_1^p \mathbf{a}_k \cdot u_k\ R(A_k) = y,\ u_k \in \mathbb{R}^d \quad \text{for } 1 \le k \le p\Big\}.
\end{aligned}
$$

Finally the LDP holds for $\langle L_n, \mathbf{f}\rangle$ by (5.2).
• We now establish the following equality:

$$
I_{\mathbf{f}}(y) = \inf_{u \in L_d^1}\{\int_{\mathcal{X}} \Lambda^*(u(x))\, R(dx),\ \int_{\mathcal{X}} \mathbf{f}(x) \cdot u(x)\, R(dx) = y\}.
\tag{5.3}
$$

Suppose that u_ϵ is an ϵ-minimizer, *i.e.* $\int_{\mathcal{X}} \Lambda^*(u_\epsilon(x))\, R(dx) \le I_{\mathbf{f}}(y) + \epsilon$ and $\int \mathbf{f} \cdot u_\epsilon\, dR = y$. Consider $u_k = \int_{A_k} u_\epsilon(x)\, R(dx)/R(A_k)$ then $\sum \mathbf{a}_k \cdot u_k\ R(A_k) = y$ and

$$\sum_1^p R(A_k)\ \Lambda^*(u_k) \overset{\text{(Jensen)}}{\le} \sum_1^p \int_{A_k} \Lambda^*(u_\epsilon(x)) \frac{R(dx)}{R(A_k)} R(A_k) = \int_{\mathcal{X}} \Lambda^*(u_\epsilon(x))\, R(dx).$$

Thus $I_{\mathbf{f}}(y) \le \inf\{\int \Lambda^*(u)\, dR,\ \int \mathbf{f} \cdot u\, dR = y\}$. The converse inequality is straightforward. Hence (5.3) is proved and so is Lemma 5.2. □

Lemma 5.3 *Assume that (A-1), (A-2) and (A-3) hold and let* $\mathbf{f} : \mathcal{X} \to \mathbb{R}^{m\times d}$ *be a bounded continuous function. Then there exists a sequence* $(\mathbf{f}^p)_{p\ge 1}$ *of finite-range step functions satisfying assumptions of Lemma 5.2 and such that* $(\langle L_n, \mathbf{f}^p\rangle)_{p\ge 0}$ *is an exponential approximation of* $\langle L_n, \mathbf{f}\rangle$, *i.e.:*

$$\lim_{p\to\infty} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(|\langle L_n, \mathbf{f}^p\rangle - \langle L_n, \mathbf{f}\rangle| > \delta) = -\infty \quad \text{for all} \quad \delta > 0.$$

Moreover, $(\langle L_n, \mathbf{f}\rangle)$ *satisfies the LDP with the good rate function*

$$\Upsilon(y) = \sup_{\epsilon>0} \liminf_{p\to\infty} \inf_{y'\in B(y,\epsilon)} I_{\mathbf{f}^p}(y') \quad \text{for } y \in \mathbb{R}^m,$$

where $B(y,\epsilon) = \{y' \in \mathbb{R}^m, |y' - y| < \epsilon\}$.

Proof [Proof of Lemma 5.3] • *Approximation of* $\mathbf{f}$ *by "good" step functions.*
As $\mathbf{f}$ is continuous and bounded, $\mathbf{f}(\mathcal{X})$ is relatively compact in $\mathbb{R}^{m\times d}$. Hence, by Proposition A.1 there exist $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_p \in \mathbb{R}^{m\times d}$ and $\epsilon_1, \dots, \epsilon_p \le \epsilon$ such that

$$\mathcal{X} \subset \cup_{k=1}^p \mathbf{f}^{-1} B(\boldsymbol{\alpha}_k, \epsilon_k) \quad \text{where} \quad R(\partial \mathbf{f}^{-1} B(\boldsymbol{\alpha}_k, \epsilon_k)) = 0,$$

where $B(\boldsymbol{\alpha}_k, \epsilon_k)$ is an open ball centered in $\boldsymbol{\alpha}_k$ with radius ϵ_k. Since each set $\mathbf{f}^{-1}B(\boldsymbol{\alpha}_k, \epsilon_k)$ is open, Assumption (A-2) yields that it is either empty or with strictly positive R-measure. Let us keep the ones with strictly positive R-measure. Hence, we get a cover of $\mathcal{X}$ by R-continuity sets with strictly positive R-measure. Thus assumptions of the Partition lemma (Lemma A.2) are fulfilled and there exists a partition $(C_l)_{1\le l\le q}$ of $\mathcal{X}$ where each C_l is a R-continuity set and has strictly positive R-measure (Lemma A.2). In particular, for each l, there exists a k such that

$$C_l \subset \overline{\mathbf{f}^{-1} B(\boldsymbol{\alpha}_k, \epsilon_k)}. \tag{5.4}$$

Consider a pairing which associates to each l a single $k(l)$ such that (5.4) is satisfied. Denote by

$$\mathbf{f}^\epsilon = \sum_{l=1}^q \boldsymbol{\alpha}_{k(l)}\ 1_{C_l}.$$

It is then straightforward to check that $\|\mathbf{f}^\epsilon - \mathbf{f}\| \le \epsilon$. Moreover, $\mathbf{f}^\epsilon$ satisfies the properties stated in Lemma 5.2. In particular, $\mathbf{f}^\epsilon$ satisfies the LDP with good rate function $I_{\mathbf{f}^\epsilon}$. Now let $p = [\epsilon^{-1}]$ where $[\cdot]$ denotes the integer part and consider the associated step function

$\mathbf{f}^p$ to obtain the uniformly convergent sequence.

• *The exponentially good approximation and the weak LDP.* As previously, consider $\mathbf{f} : \mathcal{X} \to \mathbb{R}^{m\times d}$ and let $\mathbf{f}^p$ be the approximating step function as built before. Let η be fixed and consider $\{|\langle L_n, \mathbf{f}\rangle - \langle L_n, \mathbf{f}^p\rangle| > \eta\}$. Then there exists p_ϵ such that for $p \geq p_\epsilon$, $\|\mathbf{f}^p - \mathbf{f}\| < \epsilon$. Hence

$$\{|\langle L_n, \mathbf{f}\rangle - \langle L_n, \mathbf{f}^p\rangle| > \eta\} = \{|1/n\sum_1^n(\mathbf{f} - \mathbf{f}^p)\cdot Z_i| > \eta\} \subset \{1/n\sum_1^n \epsilon|Z_i| > \eta\}.$$

Thus

$$\begin{aligned}&\limsup_{n\geq 1}\frac{1}{n}\log\mathbb{P}\{|\langle L_n, \mathbf{f}\rangle - \langle L_n, \mathbf{f}^p\rangle| > \eta\}\\ &\qquad\qquad \leq \limsup_{n\geq 1}\frac{1}{n}\log\mathbb{P}\left\{1/n\sum_1^n\epsilon|Z_i| > \eta\right\} \leq -\Lambda^*_{|Z|}\Big(\frac{\eta}{\epsilon}\Big).\end{aligned}$$

By (A-3), $\lim_{\epsilon\to 0}\Lambda^*_{|Z|}(\eta/\epsilon) = +\infty$ which yields

$$\limsup_{p\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\{|\langle L_n, \mathbf{f}\rangle - \langle L_n, \mathbf{f}^p\rangle| > \eta\} = -\infty.$$

Thus, $\langle L_n, \mathbf{f}^p\rangle$ is an exponentially good approximation of $\langle L_n, \mathbf{f}\rangle$. As each $\mathbf{f}^p$ satisfies assumptions of Lemma 5.2, $\langle L_n, \mathbf{f}^p\rangle$ satisfies the LDP and by Theorem 4.2.16 in [9], $\langle L_n, \mathbf{f}\rangle$ satisfies a weak LDP with the rate function

$$\Upsilon(y) = \sup_{\epsilon>0}\liminf_{p\to\infty}\inf_{y'\in B(y,\epsilon)} I_{\mathbf{f}^p}(y') \quad \text{for } y\in\mathbb{R}^m,$$

where $I_{\mathbf{f}^p}$ is given by (5.3).

• *Exponential tightness and the full LDP.* It remains to show that $\langle L_n, \mathbf{f}\rangle$ is exponentially tight. This is straightforward by the following inequality:

$$\begin{aligned}&\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\{|\langle L_n, \mathbf{f}\rangle| > K\}\\ &\qquad \leq \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\{1/n\sum_1^n|Z_i| > \frac{K}{\|\mathbf{f}\|}\} \leq -\Lambda^*_{|Z|}\Big(\frac{K}{\|\mathbf{f}\|}\Big) \xrightarrow[K\to\infty]{} -\infty.\end{aligned}$$

Hence the full LDP holds for $(\langle L_n, \mathbf{f}\rangle)_{n\geq 1}$ and Υ is a good rate function. Proof of Lemma 5.3 is completed. □

5.2 Identification of the rate function

Recall that

$$\begin{aligned}I_{\mathbf{f}}(y) &= \inf_{u\in L^1_d}\Big\{\int_{\mathcal{X}}\Lambda^*(u(x))\,R(dx),\ \int\mathbf{f}\cdot u\,dR = y\Big\} \quad \text{for } y\in\mathbb{R}^m,\\ \Upsilon(y) &= \sup_{\epsilon>0}\liminf_{p\to\infty}\inf_{y'\in B(y,\epsilon)} I_{\mathbf{f}^p}(y') \quad \text{for } y\in\mathbb{R}^m.\end{aligned}$$

In the case where $\mathbf{f}^p$ is a step function satisfying the assumptions of Lemma 5.2, $I_{\mathbf{f}^p}$ is lower semicontinuous as a rate function. This property is not clear for $I_{\mathbf{f}}$ if $\mathbf{f} \in C_d(\mathcal{X})$. We first clarify the link between $I_{\mathbf{f}}$ and Υ in Lemma 5.4. This is a key point to obtain the dual identity

$$\Upsilon(y) = \sup_{\theta \in \mathbb{R}^m} \{\theta \cdot y - \int_{\mathcal{X}} \Lambda(\sum_{j=1}^m \theta_j f_j(x))\, R(dx)\},$$

which is proved in Lemma 5.6.

Lemma 5.4 *Let $\mathbf{f} \in C_d(\mathcal{X})$ and assume that $(\mathbf{f}^p)_{p\geq 1}$ satisfies the properties stated in Lemma 5.3. Then Υ is the lower semicontinuous regularization of $I_{\mathbf{f}}$ (denoted in the sequel by* lscr $I_{\mathbf{f}}$*).*

The following control will be of help.

Proposition 5.5 *Let $u : \mathcal{X} \to \mathbb{R}$ be a measurable function satisfying*

$$\int_{\mathcal{X}} \Lambda^*[u(x)]\, R(dx) \leq M + 1,$$

then $\int |u|\, dR \leq K_M < \infty$ where K_M depends on M but not on u.

Proof [Proof of Proposition 5.5] By (A-3), there exists $\epsilon > 0$ such that $\{\lambda \in \mathbb{R}^d,\ |\lambda| = \epsilon\}$ is a subset of the interior of $\mathcal{D}_\Lambda$ and such that $\sup_{|\lambda|=\epsilon} \Lambda(\lambda)$ is finite. Therefore,

$$\Lambda^*(z) + \sup_{|\lambda|=\epsilon} \Lambda(\lambda) \geq \epsilon |z| \quad \text{for all } z \in \mathbb{R}^d.$$

The result follows by integrating both parts of the inequality. $\square$

Proof [Proof of Lemma 5.4] Let $u(x)$ satisfy $\int \mathbf{f} \cdot u\, dR = z$. We shall call it a control. Let us denote by

$$\begin{aligned} \mathcal{C}_p(y,\epsilon) &\triangleq \left\{u \in L^1_d(\mathcal{X}),\ |\int \mathbf{f}^p \cdot u\, dR - y| \leq \epsilon\right\}, \quad \mathcal{C}_p(y) \triangleq \mathcal{C}_p(y,0), \\ \mathcal{C}(y,\epsilon) &\triangleq \left\{u \in L^1_d(\mathcal{X}),\ |\int \mathbf{f} \cdot u\, dR - y| \leq \epsilon\right\}, \quad \mathcal{C}(y) \triangleq \mathcal{C}(y,0). \end{aligned}$$

- We first show that

$$I_{\mathbf{f}}(y) < \infty \quad \Rightarrow \quad \Upsilon(y) \leq I_{\mathbf{f}}(y). \tag{5.5}$$

Let $I_{\mathbf{f}}(y) = M < \infty$ then by Proposition 5.5,

$$I_{\mathbf{f}}(y) = \inf_{u \in L^1_d} \{\int \Lambda^*(u)\, dR,\ \int \mathbf{f} \cdot u\, dR = y,\ \int |u|\, dR \leq K_M\}.$$

Let $\epsilon > 0$ be fixed. There exists p^*_ϵ such that $\|\mathbf{f}^p - \mathbf{f}\| \leq \epsilon$ for $p \geq p^*_\epsilon$. Therefore for $p \geq p^*_\epsilon$ and under the condition $\int |u|\, dR < K_M$, one gets:

$$\mathcal{C}(y) \subset \mathcal{C}_p(y, \epsilon K_M).$$

Thus for $p \geq p_\epsilon^*$,

$$\inf_{y' \in B(y,\epsilon K_M)} \left\{ \int \Lambda^*(u)\, dR,\ u \in \mathcal{C}_p(y') \right\} \leq I_{\mathbf{f}}(y).$$

And for every $\epsilon > 0$,

$$\liminf_{p\to\infty} \inf_{y' \in B(y,\epsilon K_M)} \left\{ \int \Lambda^*(u)\, dR,\ \int \mathbf{f}^p \cdot u\, dR = y' \right\} \leq I_{\mathbf{f}}(y).$$

Finally,

$$\Upsilon(y) = \sup_{\epsilon>0} \liminf_{p\to\infty} \inf_{y' \in B(y,\epsilon K_M)} \left\{ \int \Lambda^*(u)\, dR,\ \int \mathbf{f}^p \cdot u\, dR = y' \right\} \leq I_{\mathbf{f}}(y),$$

and (5.5) is proved.
• Let us show now that

$$\Upsilon(y) < \infty \quad \Rightarrow \quad \mathrm{lscr}\, I_{\mathbf{f}}(y) \leq \Upsilon(y). \tag{5.6}$$

Let $\Upsilon(y) = M < \infty$ then by Proposition 5.5,

$$\Upsilon(y) = \sup_{\epsilon>0} \liminf_{p\to\infty} \inf_{y' \in B(y,\epsilon)} \{ \int \Lambda^*(u)\, dR,\ u \in \mathcal{C}_p(y'),\ \int |u|\, dR \leq K_M \}.$$

Let $\epsilon > 0$ be fixed. There exists p_ϵ^* such that $\|\mathbf{f}^p - \mathbf{f}\| \leq \epsilon$ for $p \geq p_\epsilon^*$. Therefore for $p \geq p_\epsilon^*$ and under the condition $\int |u|\, dR < K_M$, one gets:

$$\mathcal{C}_p(y,\epsilon) \subset \mathcal{C}(y, \epsilon(1+K_M)).$$

Thus,

$$\inf\{ \int \Lambda^*(u)\, dR,\ u \in \mathcal{C}(y,\epsilon(1+K_M))\} \leq \inf\{ \int \Lambda^*(u)\, dR,\ u \in \mathcal{C}_p(y,\epsilon)\}.$$

And for every $\epsilon > 0$,

$$\inf\{ \int \Lambda^*(u)\, dR,\ u \in \mathcal{C}(y,\epsilon(1+K_M))\} \leq \liminf_p \inf\{ \int \Lambda^*(u)\, dR,\ u \in \mathcal{C}_p(y,\epsilon)\}.$$

Finally,

$$\sup_{\epsilon'>0} \inf_{y' \in B(y,\epsilon')} I_{\mathbf{f}}(y') \leq \Upsilon(y),$$

which is the desired property since $\sup_{\epsilon'>0} \inf_{y' \in B(y,\epsilon')} I_{\mathbf{f}}(y')$ is the lower semicontinuous regularization of $I_{\mathbf{f}}$.
• Since Υ is lower semicontinuous and $\mathrm{lscr}\, I_{\mathbf{f}} \leq \Upsilon \leq I_{\mathbf{f}}$, Lemma 5.4 is proved. □

Lemma 5.6 *Let $\mathbf{f} : \mathcal{X} \to \mathbb{R}^{m\times d}$ be a bounded continuous function. Recall that*

$$I_{\mathbf{f}}(y) = \inf_{u \in L_d^1} \{ \int \Lambda^*(u)\, dR, \quad \int \mathbf{f} \cdot u\, dR = y \},$$

where $y \in \mathbb{R}^m$. Then, the following identity holds:

$$\operatorname{lscr} I_{\mathbf{f}}(z) = \sup_{\lambda \in \mathbb{R}^m} \{\lambda \cdot z - \int \Lambda(\sum_1^m \lambda_j f_j)\, dR\}.$$

In particular,

$$\Upsilon(y) = \sup_{\theta \in \mathbb{R}^m} \{\theta \cdot y - \int \Lambda(\sum_1^m \theta_j f_j)\, dR\}.$$

Proof The functionals $\int \Lambda^*\, dR$ and $\int \Lambda\, dR$ are convex conjugates for the duality (L^1_d, L^∞_d) (see Rockafellar [20, 21]), where L^∞_d denotes the space of mesurable, essentially bounded functions with value in $\mathbb{R}^d$:

$$\begin{aligned}
\int \Lambda^*(u) dR &= \sup_{g \in L^\infty_d} \{\int u \cdot g\ dR - \int \Lambda(g) dR\} \quad \text{for all } u \in L^1_d, \\
\int \Lambda(g) dR &= \sup_{u \in L^1_d} \{\int g \cdot u\ dR - \int \Lambda^*(u) dR\} \quad \text{for all } g \in L^\infty_d.
\end{aligned}$$

Recall that f_j (the j^{th} row of the matrix $\mathbf{f}$) belongs to $C_d(\mathcal{X})$. Therefore, f_j can be identified with an element of L^∞_d. Consider the operator $A : L^1_d \to \mathbb{R}^m$ and its adjoint $A^* : \mathbb{R}^m \to L^\infty_d$ defined by:

$$Au = \int \mathbf{f} \cdot u\, dR\,, \ u \in L^1_d \quad \text{and} \quad A^* y = \sum_{i=1}^m y_i f_i\,,\ y \in \mathbb{R}^m.$$

With these notations and the fact that $\int \Lambda$ and $\int \Lambda^*$ are convex conjugate, the first part of the proof is a direct application of Theorem 3 in [21]:

$$\sup_{\theta \in \mathbb{R}^m} \left\{ \theta \cdot y - \int \Lambda(A^* y)\, dR \right\} = \operatorname{lscr} \inf_{u \in L^1_d} \left\{ \int \Lambda^*(u)\, dR,\ Au = y \right\}.$$

The second part of the lemma follows from Lemma 5.4. □

5.3 Proof of Theorem 2.2

Proof Lemma 5.3 yields the LDP. Lemmas 5.4 and 5.6 yield the stated formula for the rate function. Hence Theorem 2.2 is proved. □

6 Proof of the second counterexample

Proposition 2.4 relies on Proposition 6.3 which is stated and proved in Section 6.2. We shall use the following notation:

$$\hat{Z}_n = \frac{1}{n} \sum_{i=1}^{n-1} Z_i.$$

6.1 Some preparation

We study here very carefully the rate function I which would have been associated to $\hat{Z}_n - Z_n/n$ if this quantity was to satisfy a LDP. We also establish a usefull inequality. The proofs of Propositions 6.1 and 6.2, though standard, are given for the reader's convenience. We introduce the following notations:

$$\begin{aligned}
\delta(\lambda|_{[-1,+\infty)}) &= \begin{cases} 0 & \text{if } \lambda \in [-1,\infty) \\ +\infty & \text{otherwise} \end{cases}, \\
\delta^*(z) &= \sup_{\lambda\in\mathbb{R}}\{\lambda\, z - \delta(\lambda|_{[-1,+\infty)})\} = \begin{cases} -z & \text{if } z \le 0 \\ +\infty & \text{otherwise} \end{cases}, \\
I(z) &= \inf\{\Lambda^*(x) + \delta^*(y),\ x+y=z\}.
\end{aligned}$$

Note that δ^* would have been the rate function associated to the particle $-Z_n/n$ if this particle was to satisfy a LDP. Similarily, $I(z)$ would have been the rate function associated to $\hat{Z}_n - Z_n/n$.

Proposition 6.1 *Denote by* $z_-^* = \Lambda'(-1)$.

- *If* $z \ge z_-^*$ *then* $I(z) = \Lambda^*(z)$.
- *If* $z < z_-^*$ *then* $I(z) = \Lambda^*(z_-^*) + z_-^* - z$.

In the case where $z < z_-^*$*, the infimum* $\inf\{\Lambda^*(x) + \delta^*(y),\ x+y=z\}$ *is uniquely attained for* $x = z_-^*$ *and* $y = z - z_-^*$.

Proof First note that:

$$I(z) = \sup_{\lambda\in\mathbb{R}}\{\lambda\ z - \Lambda(\lambda) - \delta(\lambda|_{[-1,+\infty)})\} = \sup_{\lambda\in[-1,1]}\{\lambda\ z - \Lambda(\lambda)\}.$$

Let us denote by $z_+^* = \Lambda'(1)$. One should notice that the special form of the probability distribution $\frac{e^{-an}}{1+a_n^3}$ implies that Λ is not steep and therefore that z_+^* is finite.
•If $z \ge z_+^*$ then

$$I(z) = \sup_{\lambda\in\mathbb{R}}\{\lambda\ (z - z_+^*) + \lambda\ z_+^* - \Lambda(\lambda)\} \overset{(a)}{=} z - z_+^* + \Lambda^*(z_+^*) \overset{(b)}{=} \Lambda^*(z),$$

where (a) and (b) are standard.
•If $z_-^* \le z \le z_+^*$ then there exists $\lambda_z \in [-1,1]$ satisfying $z = \Lambda'(\lambda_z)$ therefore $I(z) = z\lambda_z - \Lambda(\lambda_z) = \Lambda^*(z)$.
•If $z \le z_-^*$ then

$$I(z) = \sup_{\lambda\in[-1,1]}\{\lambda\ (z - z_-^*) + \lambda\ z_-^* - \Lambda(\lambda)\} = z_-^* - z + \Lambda^*(z_-^*).$$

To prove the last part of the proposition, consider the function $\Lambda^*(x) + \delta^*(z-x)$. By the second point of the proposition, this function attains its minimum for $x = z_-^*$. Since $\Lambda^*(x) + \delta^*(z-x)$ is strictly convex in a neighbourhood of z_-^*, this minimum is unique and the proposition is proved. $\square$

Proposition 6.2 *Let $\epsilon > 0$ and $z < z_-^* - 2\epsilon$ be fixed and consider the following quantities:*

$$\begin{aligned} I_\epsilon &= \inf\{\Lambda^*(x) + \delta^*(y),\ x + y \in [z - 2\,\epsilon, z + 2\,\epsilon],\ x \notin (z_-^* - \epsilon, z_-^* + \epsilon)\}, \\ I_0 &= \inf\{I(u),\ u \in [z - 2\,\epsilon, z + 2\,\epsilon]\} = \Lambda^*(z_-^*) + z_-^* - z - 2\,\epsilon. \end{aligned}$$

Then $I_\epsilon > I_0$.

Proof First note that I_0 and I_ϵ are always finite and that $I_\epsilon \geq I_0$. Assume that $I_\epsilon = I_0$. Let (x_n, y_n) be a sequence of minimizing elements, that is

$$\begin{gathered} x_n + y_n \in [z - 2\,\epsilon, z + 2\,\epsilon], \\ x_n \notin (z_-^* - \epsilon, z_-^* + \epsilon), \\ \lim_{n\to\infty} [\Lambda^*(x_n) + \delta^*(y_n)] = I_\epsilon \quad = I_0. \end{gathered}$$

Using the compactness of the level sets of Λ^* and δ^*, one can prove that there exists a minimizer (x_*, y_*) satisfying:

$$\begin{gathered} x_* + y_* \in [z - 2\,\epsilon, z + 2\,\epsilon], \\ x_* \notin (z_-^* - \epsilon, z_-^* + \epsilon), \\ \Lambda^*(x_*) + \delta^*(y_*) = I_0 = \Lambda^*(z_-^*) + z_-^* - z - 2\epsilon. \end{gathered}$$

The second part of Proposition 6.1 yields that $x_* = z_-^*$, which contradicts $x_* \notin (z_-^* - \epsilon, z_-^* + \epsilon)$. Necessarily, $I_\epsilon > I_0$ and Proposition 6.2 is proved. □

6.2 Statement and proof of Proposition 6.3

Let $\epsilon > 0$ and $z < z_-^* - 2\epsilon$ be fixed. The following holds.

Proposition 6.3 *There exists a finite real number $A > 0$ such that*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\{\hat{Z}_n - Z_n/n \in [z - 2\,\epsilon, z + 2\,\epsilon]\} \leq -A. \tag{6.1}$$

Moreover, there exist real numbers $\alpha \in (0,1)$, $\delta \in (0, (1-\alpha)2\,\epsilon)$ and a real number $B \in (0, A)$ such that:

$$-B \leq \limsup_{n} \frac{1}{n} \log \mathbb{P}\left\{\hat{Z}_n - Z_n/n \in [z + \alpha\, 2\,\epsilon - \delta, z + \alpha\, 2\,\epsilon + \delta]\right\}. \tag{6.2}$$

The proof of Proposition 6.3, though very involved, is interesting because it gives an insight on how Large Deviation phenomena fail to occur.

Proof [Proof of Proposition 6.3] • We first prove that there exists a subsequence $\phi(n)$ such that

$$\limsup_{n\to\infty} \frac{1}{\phi(n)} \log \mathbb{P}\{\hat{Z}_{\phi(n)} - Z_{\phi(n)}/\phi(n) \in (z - 2\epsilon, z + 2\epsilon)\} \leq -I_\epsilon < -I_0, \tag{6.3}$$

where I_ϵ and I_0 are defined in Proposition 6.2. This will yield the first part of Proposition 6.3. Consider the following notations:

$$\begin{aligned}
B_k^1 &= [z_-^* + 2k\epsilon - \epsilon, z_-^* + 2k\epsilon + \epsilon), \\
B_k^2 &= (z - z_-^* - 2k\epsilon - 3\epsilon, z - z_-^* - 2k\epsilon + 3\epsilon].
\end{aligned}$$

The following is straightforward:

$$\{\hat{Z}_n - Z_n/n \in [z - 2\epsilon, z + 2\epsilon]\} \subset \bigcup_{k \in \mathbb{Z}} \left\{\hat{Z}_n \in B_k^1\right\} \cap \left\{-Z_n/n \in B_k^2\right\}.$$

The previous union is a union of disjoint sets. Moreover $\{\hat{Z}_n \in B_k^1\}$ is empty if k is negative with $|k|$ large enough (say $k \le k^- < 0$) since $\hat{Z}_n$ is a nonnegative random variable. Therefore, the following holds:

$$\mathbb{P}\{\hat{Z}_n - Z_n/n \in [z - 2\epsilon, z + 2\epsilon]\} \le \sum_{k \ge k^-} \mathbb{P}\left\{\hat{Z}_n \in B_k^1\right\} \mathbb{P}\left\{-Z_n/n \in B_k^2\right\}.$$

Let $L > I_\epsilon$ be given. By Cramér's theorem, there exists k^+ such that

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}\left\{\hat{Z}_n \ge z_-^* + k^+ \epsilon\right\} \le -L.$$

Furthermore,

$$\begin{aligned}
&\mathbb{P}\{\hat{Z}_n - Z_n/n \in [z - 2\epsilon, z + 2\epsilon]\} \\
&\qquad \le \sum_{k=k^-}^{k^+} \mathbb{P}\{\hat{Z}_n \in B_k^1\} \mathbb{P}\{-Z_n/n \in B_k^2\} \\
&\qquad\qquad + \mathbb{P}\left\{\hat{Z}_n \ge z_-^* + k^+ \epsilon\right\}. \qquad (6.4)
\end{aligned}$$

Let us give the guideline of what is done next: We know from Proposition 6.1 that the infimum

$$\inf\{I(y), y \in [z - 2\epsilon, z + 2\epsilon]\} = I(z - 2\epsilon)$$

is uniquely attained for $I(z - 2\epsilon) = \Lambda^*(z_-^*) + \delta^*(z - z_-^* - 2\epsilon)$. Therefore, consider

$$\mathcal{T} = \{\hat{Z}_n \approx z_-^*\} \cap \{-Z_n/n \approx z - z_-^* - 2\epsilon\}.$$

The event $\mathcal{T}$ can be seen as the "most typical" subset of $\{\hat{Z}_n - Z_n/n \in [z - 2\epsilon, z + 2\epsilon]\}$ since the infimum of the rate function is "realized" on $\mathcal{T}$. If we choose a subsequence $\phi(n)$ such that

$$\left\{-\frac{Z_{\phi(n)}}{\phi(n)} \approx z - z_-^* - 2\epsilon\right\} = \emptyset,$$

then the Large Deviation upper bound shall decrease and we shall obtain (6.3). We formalize this in the sequel.

Consider the subsequence defined by $\phi(n) = \left[\frac{2a_n}{z_-^* + \epsilon - z}\right]$ where $[x]$ denotes the integer part of x. It is then straightforward to check that

$$\mathbb{P}\left\{ -\frac{Z_{\phi(n)}}{\phi(n)} \in (z - z_-^* - 3\epsilon, z - z_-^* + 3\epsilon] \right\} = 0$$

for n large enough. Therefore, when considering the subsequence $\phi(n)$, (6.4) becomes

$$\begin{aligned} &\mathbb{P}\{\hat{Z}_{\phi(n)} - Z_{\phi(n)}/\phi(n) \in [z - 2\epsilon, z + 2\epsilon]\} \\ &\qquad\leq \sum_{k\neq 0, k=k^-}^{k^+} \mathbb{P}\{\hat{Z}_{\phi(n)} \in B_k^1\}\mathbb{P}\{-Z_{\phi(n)}/\phi(n) \in B_k^2\} \\ &\qquad\qquad + \mathbb{P}\left\{\hat{Z}_{\phi(n)} \geq z_-^* + k^+\epsilon\right\}. \end{aligned}$$

In other words, the "most typical" subset

$$\left\{ -\frac{Z_{\phi(n)}}{\phi(n)} \in (z - z_-^* - 3\epsilon, z - z_-^* + 3\epsilon] \right\} \cap \{\hat{Z}_{\phi(n)} \in [z_-^* - \epsilon, z_-^* + \epsilon)\}$$

has been removed. Let $k \in \{k^-, \dots, k^+\}$ and $k \neq 0$. Usual techniques to derive upper bounds yield:

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{\phi(n)} \log \mathbb{P}\{\hat{Z}_{\phi(n)} \in B_k^1\}\mathbb{P}\{-Z_{\phi(n)}/\phi(n) \in B_k^2\} \\ &\quad\leq \; -\inf\{\Lambda^*(x),\ x \in B_k^1\} - \inf\{\delta^*(y),\ y \in B_k^2\} \\ &\quad\leq \; -\inf\{\Lambda^*(x) + \delta^*(y),\ x + y \in [z - 2\epsilon, z + 2\epsilon],\ x \in B_k^1\} \end{aligned}$$

and by Lemma 1.2.15 in [9],

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{\phi(n)} \log \mathbb{P}\{\hat{Z}_{\phi(n)} - Z_{\phi(n)}/\phi(n) \in [z - 2\epsilon, z + 2\epsilon]\} \\ &\quad\leq \sup_{k\neq 0,\ k^-\leq k\leq k^+} -\inf\{\Lambda^*(x) + \delta^*(y),\ x + y \in [z - 2\epsilon, z + 2\epsilon],\ x \in B_k^1\} \cdots \\ &\qquad\qquad \vee (-L) \\ &\quad\leq \; -\inf\{\Lambda^*(x) + \delta^*(y),\ x + y \in [z - 2\epsilon, z + 2\epsilon],\ x \notin (z_-^* - \epsilon, z_-^* + \epsilon)\} \\ &\quad\leq \; -I_\epsilon < -I_0, \end{aligned}$$

where $a \vee b = \sup(a, b)$ and where the last inequality comes from Proposition 6.2. The first part of the proposition is then proved.

• Let us now prove the second part of the proposition. Consider $\{\hat{Z}_n - Z_n/n \in [z + \alpha 2\epsilon - \delta, z + \alpha 2\epsilon + \delta]\}$ where $\alpha < 1$ and $\delta < (1-\alpha)2\,\epsilon$. Then

$$\begin{aligned} &\{\hat{Z}_n \in [z_-^* - \delta/2, z_-^* + \delta/2]\} \cap \{-Z_n/n \in [z - z_-^* + \alpha 2\epsilon - \delta/2, z - z_-^* + \alpha 2\epsilon + \delta/2]\} \\ &\qquad\qquad \subset \{\hat{Z}_n - Z_n/n \in [z + \alpha 2\epsilon - \delta, z + \alpha 2\epsilon + \delta]\}. \end{aligned}$$

Choose now a subsequence of integers defined by the following inequalities:

$$\begin{cases} a_{n-1} & < \phi(n)(z_-^* - z - \alpha\, 2\,\epsilon - \delta/2) & \leq a_n \\ a_n & \overset{(a)}{\leq} \phi(n)(z_-^* - z - \alpha\, 2\,\epsilon + \delta/2) & < a_{n+1} \end{cases},$$

where $a_n = 16^n$. Such a subsequence always exists and

$$\liminf_{n\to\infty} \frac{1}{\phi(n)} \log \mathbb{P}\left\{-\frac{Z_{\phi(n)}}{\phi(n)} \in [z - z_-^* + \alpha 2\epsilon - \delta/2, z - z_-^* + \alpha 2\,\epsilon + \delta/2]\right\}$$
$$= \liminf_{n\to\infty} \frac{1}{\phi(n)} \log \mathbb{P}\{Z = a_n\} = \liminf_{n\to\infty} -\frac{a_n}{\phi(n)} \geq -(z_-^* - z - \alpha\, 2\,\epsilon + \delta/2),$$

where the last inequality comes from (a). As by Cramér's theorem,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\{\hat{Z}_n \in [z_-^* - \delta/2, z_-^* + \delta/2]\} \geq -\Lambda^*(z_-^*),$$

we obtain:

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\{\hat{Z}_n - Z_n/n \in [z + \alpha 2\epsilon - \delta, z + \alpha 2\epsilon + \delta]\} \geq -(\Lambda^*(z_-^*) + z_-^* - z - \alpha\, 2\,\epsilon + \delta/2).$$

Finally, as $\lim_{\alpha\to 1, \delta<(1-\alpha)\, 2\,\epsilon}(\Lambda^*(z_-^*) + z_-^* - z - \alpha\, 2\,\epsilon + \delta/2) = I_0$ and $I_0 < I_\epsilon$, there exist $\alpha > 0$, $\delta > 0$ and $\gamma > 0$ such that $\Lambda^*(z_-^*) + z_-^* - z - \alpha\, 2\,\epsilon + \delta/2 \leq I_\epsilon - \gamma$ and the second part of the proposition is proved. □

6.3 End of proof

We prove here Proposition 2.4.

Proof [Proof of Proposition 2.4] Assume that $(\hat{Z}_n - Z_n/n)_{n\geq 1}$ satisfies a LDP with rate function J and choose the real numbers α and δ as in Proposition 6.3. Then:

$$-\inf_{z\in(z-2\,\epsilon, z+2\,\epsilon)} J(z) \leq \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\{\hat{Z}_n - Z_n/n \in (z - 2\epsilon, z + 2\epsilon)\} \overset{(a)}{\leq} -A,$$

where (a) follows from (6.1) and

$$-B \overset{(b)}{\leq} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left\{\hat{Z}_n - Z_n/n \in [z + \alpha\, 2\epsilon - \delta, z + \alpha\, 2\epsilon + \delta]\right\}$$
$$\leq -\inf_{[z+\alpha\, 2\epsilon-\delta, z+\alpha\, 2\epsilon+\delta]} J(z),$$

where (b) follows from (6.2). As $B < A$, one should have

$$\inf_{u\in[z+\alpha\, 2\epsilon-\delta, z+\alpha\, 2\epsilon+\delta]} J(u) < \inf_{u\in(z-2\epsilon, z+2\epsilon)} J(u),$$

which is impossible since $[z + \alpha\, 2\epsilon - \delta, z + \alpha\, 2\epsilon + \delta] \subset (z - 2\epsilon, z + 2\epsilon)$. □

7 Proof of Proposition 3.5

Proof One can easily check that the measure defined by $d\gamma = E\,d\ell$ is always a minimizer of I under $\mathcal{C}$ (Jensen). Therefore, if μ is a minimizer:

$$I(\mu) = \Lambda^*(E) = \begin{cases} \Lambda^*(E) & \text{if } E < z^* \\ \Lambda^*(z^*) + E - z^* & \text{else.} \end{cases}.$$

In the case where $E < z^*$, let us prove that the minimizer is unique. Assume that μ and ν are distinct minimizers, that is:

$$\begin{cases} \langle \mu, 1 \rangle = E \\ \langle \nu, 1 \rangle = E \end{cases} \quad \text{and} \quad I(\mu) = I(\nu) = \Lambda^*(E).$$

Note that any convex combination $\alpha\mu + \beta\nu$ belongs to $\mathcal{C}$.

$$I(\alpha\mu + \beta\nu) = \int \Lambda^*\left(\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell}\right) d\ell + \alpha\mu_s[0,1] + \beta\nu_s[0,1].$$

If $\ell\left\{\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell} < z^*\right\} = 0$ then

$$I(\alpha\mu + \beta\nu) \geq \int_{\left\{\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell} \geq z^*\right\}} \Lambda^*\left(\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell}\right) d\ell \geq \Lambda^*(z^*) > \Lambda^*(E),$$

which is impossible since $I(\alpha\mu + \beta\nu) \leq \alpha I(\mu) + \beta I(\nu) = \Lambda^*(E)$. Therefore, assume that $\ell\left\{\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell} < z^*\right\} > 0$. In this case,

$$\begin{aligned} &\int \Lambda^*\left(\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell}\right) d\ell \\ &= \int_{\left\{\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell} < z^*\right\}} \Lambda^*\left(\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell}\right) d\ell + \int_{\left\{\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell} \geq z^*\right\}} \Lambda^*\left(\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell}\right) d\ell. \end{aligned}$$

Since Λ^* is strictly convex for $z \in (-\infty, z^*]$,

$$\Lambda^*\left(\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell}\right) < \alpha\Lambda^*\left(\frac{d\mu_a}{d\ell}\right) + \beta\Lambda^*\left(\frac{d\nu_a}{d\ell}\right)$$

on $\left\{\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell} < z^*\right\}$. Moreover, the inequality remains strict by integrating (7) over the set $\left\{\alpha\frac{d\mu_a}{d\ell} + \beta\frac{d\nu_a}{d\ell} < z^*\right\}$. Therefore, $I(\alpha\mu + \beta\nu) < \alpha I(\mu) + \beta I(\nu) = \Lambda^*(E)$, which is impossible. Necessarily, the minimizer is unique.

In the case where $c \geq z^*$, the only thing to prove is that $\ell\left\{\frac{d\mu_a}{d\ell} < z^*\right\} = 0$ if μ is a minimizer. Assume that $\ell\left\{\frac{d\mu_a}{d\ell} < z^*\right\} > 0$. We denote by $g(x) = \frac{d\mu_a}{dx}(x)$.

$$\int \Lambda^*(g)\,d\ell = \int_{\{g < z^*\}} \Lambda^*(g)\,d\ell + \int_{\{g \geq z^*\}} \Lambda^*(g)\,d\ell.$$

Jensen's inequality yields

$$\int_{\{g<z^*\}} \Lambda^*(g)\,d\ell \quad \geq \quad \ell\{g<z^*\}\Lambda^*\left(\frac{\int_{\{g<z^*\}} g\,d\ell}{\ell\{g<z^*\}}\right).$$

Similarily,

$$\int_{\{g\geq z^*\}} \Lambda^*(g)\,d\ell \quad \geq \quad \ell\{g\geq z^*\}\Lambda^*\left(\frac{\int_{\{g\geq z^*\}} g\,d\ell}{\ell\{g\geq z^*\}}\right).$$

Therefore,

$$\int \Lambda^*(g)\,d\ell \quad \geq \quad \ell\{g<z^*\}\Lambda^*\left(\frac{\int_{\{g<z^*\}} g\,d\ell}{\ell\{g<z^*\}}\right)+\ell\{g\geq z^*\}\Lambda^*\left(\frac{\int_{\{g\geq z^*\}} g\,d\ell}{\ell\{g\geq z^*\}}\right).$$

Since $\frac{\int_{\{g<z^*\}} g\,dx}{\ell\{g<z^*\}} < z^*$ and Λ^* is strictly convex on $(-\infty, z^*]$, we have:

$$\ell\{g<z^*\}\Lambda^*\left(\frac{\int_{\{g<z^*\}} g\,d\ell}{\ell\{g<z^*\}}\right)+\ell\{g\geq z^*\}\Lambda^*\left(\frac{\int_{\{g\geq z^*\}} g\,d\ell}{\ell\{g\geq z^*\}}\right) \quad > \quad \Lambda^*\left(\int g\,d\ell\right),$$

and $I(\mu) > \Lambda^*(\langle\mu_a, 1\rangle) + \langle\mu_s, 1\rangle$. Denote by $\alpha E = \langle\mu_a, 1\rangle$ where $\alpha \in (0,1)$. As the maximun slope of the convex function Λ^* is 1, the following inequality is true for every real number E:

$$\frac{\Lambda^*(E)-\Lambda^*(\alpha E)}{E-\alpha E} \leq 1.$$

Therefore, $I(\mu) > \Lambda^*(\alpha E) + E - \alpha E \geq \Lambda^*(E)$. This is impossible since $I(\mu) = \Lambda^*(E)$ as μ is a minimizer. Necessarily, $\ell\left\{\frac{d\mu_a}{d\ell} < z^*\right\} = 0$ and the second part of the proposition is proved. □

A Two results concerning R-continuity and partitioning

It has been seen in Lemma 5.2 that step functions $\mathbf{f}(x) = \sum_{k=1}^{p} \mathbf{a}_k 1_{A_k}(x)$ where

$$\inf\{R(A_k),\ 1\leq k\leq p\} > 0 \quad \text{and} \quad \sup\{R(\partial A_k),\ 1\leq k\leq p\} = 0,$$

are relevant to get LDPs. The two following results permit us to construct step functions of the previous kind which approximate any bounded continuous function. These results are used in the proof of Lemma 5.3.

A.1 How to get R-continuity in a cover of $\mathcal{X}$?

In the following proposition, $\mathcal{X}$ is a topological space endowed with its Borel σ-field and with a probability distribution R, $\mathcal{Y}$ is a metric space endowed with its Borel σ-field. As usual, $B(y,\epsilon)$ is the ball centered at $y \in \mathcal{Y}$ with radius $\epsilon > 0$. Recall that A is a R-continuity set if

$$R(\partial A) = 0 \quad \text{where } \partial A = \bar{A} - \mathring{A}.$$

Here, $\bar{A}$ denotes the closure of A and $\mathring{A}$ its interior. Let $f : \mathcal{X} \to \mathcal{Y}$. The following inclusion will be usefull:

$$\partial f^{-1}(A) \subset f^{-1}(\partial A). \tag{A.1}$$

Proposition A.1 *Let $f : \mathcal{X} \to \mathcal{Y}$ be a continuous function. Assume moreover that the range $f(\mathcal{X})$ is relatively compact in $\mathcal{Y}$. Then for every $\epsilon > 0$, there exist $(y_k, 1 \le k \le p) \subset \mathcal{Y}$ and $\epsilon_1, \dots, \epsilon_p \in (0, \infty)$ such that $\epsilon_k \le \epsilon$ and*

$$\mathcal{X} \subset \cup_{k=1}^{p} f^{-1}B(y_k, \epsilon_k) \quad \textit{where} \quad R(\partial f^{-1}B(y_k, \epsilon_k)) = 0 \quad \textit{for } k \in \{1, \dots, p\}.$$

Proof Let $\epsilon > 0$ be fixed and denote by

$$\begin{aligned} \mathcal{I}_\epsilon(y) &= \{\epsilon' \in (0, \epsilon],\ R(\partial f^{-1}B(y, \epsilon')) > 0\}, \\ \mathcal{J}_\epsilon(y) &= \{\epsilon' \in (0, \epsilon],\ R(\partial f^{-1}B(y, \epsilon')) = 0\}. \end{aligned}$$

The set $\mathcal{I}_\epsilon(y)$ is at most countable. In fact, consider

$$\begin{aligned} \phi : (0, \epsilon] &\to [0, 1] \\ \epsilon' &\mapsto R \circ f^{-1}\ B(y, \epsilon'). \end{aligned}$$

The function ϕ is non-decreasing, left-countinuous and admits at most a countable number of discontinuities. Let us show that if ϕ is continuous in ϵ_0, then

$$R \circ f^{-1}\,[\partial B(y, \epsilon_0)] = 0. \tag{A.2}$$

In fact,

$$\lim_{\epsilon' \searrow \epsilon_0} R \circ f^{-1}B(y, \epsilon') = R \circ f^{-1}B(y, \epsilon_0)$$

by continuity. But

$$\begin{aligned} \lim_{\epsilon' \searrow \epsilon_0} R \circ f^{-1}B(y, \epsilon') &= R \circ f^{-1} \cap_{\epsilon' > \epsilon} B(y, \epsilon') \\ &= R \circ f^{-1}\bar{B}(y, \epsilon_0), \end{aligned}$$

where $\bar{B}(y, \epsilon_0)$ denotes the closed ball centered in y and with radius ϵ_0. Finally,

$$R \circ f^{-1}\,[\partial B(y, \epsilon_0)] = R \circ f^{-1}\bar{B}(y, \epsilon_0) - R \circ f^{-1}B(y, \epsilon_0) = 0.$$

The inclusion (A.1) yields that

$$R\partial f^{-1}B(y, \epsilon_0) \le R \circ f^{-1}\,[\partial B(y, \epsilon_0)] = 0.$$

Thus, $\mathcal{I}_\epsilon(y)$ is at most countable. As $\mathcal{I}_\epsilon(y) \cup \mathcal{J}_\epsilon(y) = (0, \epsilon]$, $\mathcal{J}_\epsilon(y)$ is never empty and

$$f(\mathcal{X}) \subset \cup_{y \in f(\mathcal{X}),\ \epsilon' \in \mathcal{J}_\epsilon(y)} B(y, \epsilon').$$

As $f(\mathcal{X})$ is relatively compact, there exist $(y_k, 1 \le k \le p) \subset \mathcal{Y}$ and $(\epsilon_k, 1 \le k \le p) \subset (0, \infty)$ satisfying:

$$f(\mathcal{X}) \subset \cup_{k=1}^{p} B(y_k, \epsilon_k) \quad \Rightarrow \quad \mathcal{X} \subset \cup_{k=1}^{p} f^{-1}B(y_k, \epsilon_k).$$

Moreover, as $\epsilon_k \in \mathcal{J}_\epsilon(y_k)$, $R(\partial f^{-1}B(y_k, \epsilon_k)) = 0$ and Proposition A.1 is proved. $\square$

A.2 A Partition Lemma

Lemma A.2 *Assume that (A-2) holds and let* $(A_k, 1 \leq k \leq p)$ *be a measurable cover of* $\mathcal{X}$ *satisfying* $R(A_1) > 0$ *and* $R(\partial A_k) = 0$ *for* $1 \leq k \leq p$. *Then there exists a partition* $(B_l, 1 \leq l \leq q)$ *of* $\mathcal{X}$ *satisfying* $B_1 = \bar{A}_1$, $R(B_l) > 0$ *and* $R(\partial B_l) = 0$ *for* $1 \leq l \leq q$. *Moreover, for each* B_l, *there exists* A_k *such that* $B_l \subset \bar{A}_k$.

Proof We proceed by induction on p. If $p = 1$ we take $B_1 = \bar{A}_1$ and the result is proved. Let $p > 1$. We can assume that the cover is based on closed sets. In fact, if $(A_k)_{1 \leq k \leq p}$ is a cover satisfying the assumptions of the lemma, so is $(\bar{A}_k)_{1 \leq k \leq p}$. If $\mathcal{X} \subset \bar{A}_1$ then the partition is reduced to the single element $B_1 = \bar{A}_1$. Otherwise $\mathcal{X} \backslash \bar{A}_1$ is a nonempty open set and $R(\mathcal{X} \backslash \bar{A}_1) > 0$ by (A-2). Necessarily there exists k satisfying $R(\bar{A}_k \backslash \bar{A}_1) > 0$. In fact

$$0 < R(\mathcal{X} \backslash \bar{A}_1) \leq \sum_2^p R(\bar{A}_k \backslash \bar{A}_1).$$

Now consider the family $\{\bar{A}_1 \cup \bar{A}_k,\ A_j,\ 2 \leq j \leq p, j \neq k\}$. This is a cover of $p-1$ elements of $\mathcal{X}$ satisfying the assumptions of the lemma (recall that the R-continuity sets form an algebra on $\mathcal{X}$). Hence we can apply the induction assumption and there exists a partition $(B_l, 1 \leq l \leq q)$ where $B_1 = \bar{A}_1 \cup \bar{A}_k$, $R(\partial B_l) = 0$ and $R(B_l) > 0$ for $1 \leq l \leq q$. Now split B_1 into $C_1 = \bar{A}_1$ and $C_2 = \bar{A}_k \backslash \bar{A}_1$ then the partition $\{C_1,\ C_2,\ B_l,\ 2 \leq l \leq q\}$ satisfies the requirements of the lemma and Lemma A.2 is proved. □

Acknowledgments

I am indebted to Professor Christian Léonard for his constant support during the preparation of this work. I also wish to thank Professor Amir Dembo for many valuable comments.

References

[1] G. Ben Arous, A. Dembo, and A. Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1):1–67, 2001.

[2] R.R. Bahadur and S. Zabell. Large deviations of the sample means in general vector spaces. *Ann. Probab.*, 7:587–621, 1979.

[3] B. Bercu, F. Gamboa, and M. Lavielle. Sharp large deviations for gaussian quadratic forms with applications. *ESAIM Probab. Statist.*, 4:1–24, 2000.

[4] B. Bercu, F. Gamboa, and A. Rouault. Large deviations for quadratic functionals of stationary Gaussian processes. *Stochastic Process. Appl.*, 71:75–90, 1997.

[5] W. Bryc and A. Dembo. Large deviations for quadratic functionals of Gaussian processes. *J. Theoret. Probab.*, 10:307–332, 1997.

[6] I. Csiszár, F. Gamboa, and E. Gassiat. Mem pixel correlated solutions for generalized moment and interpolation problems. *IEEE Trans. Inform. Theory*, 45(7):2253–2270, 1999.

[7] D. Dacunha-Castelle and F. Gamboa. Maximum d'entropie et problème des moments. *Ann. Inst. H. Poincaré Probab. Statist.*, 26:567–596, 1990.

[8] A. de Acosta. Large deviations for vector-valued Lévy processes. *Stochastic Process. Appl.*, 51:75–115, 1994.

[9] A. Dembo and O. Zeitouni. *Large Deviations Techniques And Applications.* Springer Verlag, New York, second edition, 1998.

[10] N. Dunford and J.T. Schwartz. *Linear Operators, Part I.* Interscience Publishers Inc., New York, 1958.

[11] R.S. Ellis, J. Gough, and J.V. Pulé. The large deviation principle for measures with random weights. *Rev. Math. Phys.*, 5(4):659–692, 1993.

[12] F. Gamboa and E. Gassiat. Bayesian methods and maximum entropy for ill-posed inverse problems. *Ann. Statist.*, 25(1):328–350, 1997.

[13] F. Gamboa, A. Rouault, and M. Zani. A functional large deviations principle for quadratic forms of Gaussian stationary processes. *Statist. Probab. Lett.*, 43:299–308, 1999.

[14] C. Léonard. Large deviations for Poisson random measures and processes with independent increments. *Stochastic Process. Appl.*, 85:93–121, 2000.

[15] C. Léonard and J. Najim. An extention of Sanov's theorem. Application to the Gibbs conditionning principle. submitted, 2000.

[16] J. Lynch and J. Sethuraman. Large deviations for processes with independent increments. *Ann. Probab.*, 15(2):610–627, 1987.

[17] A.A. Mogul'skii. Large deviations for trajectories of multi-dimentional random walks. *Theory Probab. Appl.*, 21:300–315, 1976.

[18] A.A. Mogul'skii. Large deviations for processes with independent increments. *Ann. Probab.*, 21:202–213, 1993.

[19] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, 1970.

[20] R. T. Rockafellar. Convex integral functionals and duality. In E. Zarantello, editor, *Contributions to Non Linear Functional Analysis*, pages 215–236. Academic Press, 1971.

[21] R. T. Rockafellar. Integrals which are convex functionals, II. *Pacific J. Math.*, 39(2):439–469, 1971.

[22] M. van den Berg, T. C. Dorlas, J. T. Lewis, and J. V. Pulé. A perturbed mean field model of an interacting boson gas and the large deviation principle. *Comm. Math. Phys.*, 127(1):41–69, 1990.

[23] M. Zani. *Grandes déviations pour des fonctionnelles issues de la statistique des processus.* PhD thesis, Université Paris-Sud, 1999.