The Annals of Applied Probability
2002, Vol. 12, No. 2, 607–636

BINS AND BALLS: LARGE DEVIATIONS OF THE EMPIRICAL OCCUPANCY PROCESS

BY STÉPHANE BOUCHERON, FABRICE GAMBOA AND CHRISTIAN LÉONARD

Université Paris-Sud, Université Toulouse III and Université Paris X

In the random allocation model, balls are sequentially inserted at random into n exchangeable bins. The *occupancy score* of a bin denotes the number of balls inserted in this bin. The (random) distribution of occupancy scores defines the object of this paper: the *empirical occupancy measure* which is a probability measure over the integers. This measure-valued random variable packages many useful statistics. This paper characterizes the *large deviations* of the flow of empirical occupancy measures when n goes to infinity while the number of inserted balls remains proportional to n. The main result is a Sanov-like theorem for the empirical occupancy measure when the set of probability measures over the integers is endowed with metrics that are slightly stronger than the total variation distance. Thanks to a coupling argument, this result applies to the degree distribution of sparse random graphs.

1. Introduction. Consider the following classical model in random combinatorics. At each time $k = 1, 2, \dots,$ a ball is thrown into one bin among n. Let $\{1, \dots, n\}$ denote the set of bins. The set $\Omega^n = \{1, \dots, n\}^{\{1,2,\dots\}}$ of all sequences in $\{1, \dots, n\}$ is the natural space for the realizations of this experiment. For any $k \geq 1$, the canonical projection $B_k^n\colon\ \omega = (\omega_l)_{l\geq 1} \in \Omega^n \mapsto \omega_k \in \{1, \dots, n\}$ is the random variable: "name of the bin into which the kth ball is thrown."

To make things easier, it is assumed that at time $k = 0$, all bins are empty. The score of bin α at time $k \geq 0$ is defined by

$$S_k^n(\alpha) = \sum_{l=1}^{k} \mathbb{1}_{\{B_l^n = \alpha\}},$$

with the convention that $S_0^n \equiv 0$. Let us consider the time-scaling $k = \lfloor nt \rfloor$, $0 \leq t \leq T$ where $\lfloor s \rfloor$ is the integer part of s. We are interested in the time-rescaled evolution of the joint empirical distribution of the scores. This is described by the following empirical occupancy process from $[0, T]$ to the set $\mathcal{P}(\mathbb{N})$ of all probability measures on $\mathbb{N}$:

$$X_t^n = \frac{1}{n} \sum_{\alpha=1}^{n} \delta_{S_{\lfloor nt \rfloor}^n(\alpha)} = \sum_{i\geq 0} X_t^n(i)\delta_i \in \mathcal{P}(\mathbb{N}), \qquad 0 \leq t \leq T,$$

Received June 2000; revised September 2001.

AMS 2000 subject classifications. Primary 60F10, 05A17, 05C40; secondary 05C80, 05A99.

Key words and phrases. Random graphs, large deviations, occupancy, Orlicz spaces, Sanov theorem, Poisson approximation.

where δ stands for the Dirac measure and $X_t^n(i)$ is the proportion of bins with score i after $\lfloor nt \rfloor$ ball allocations. The sequence (X^n) satisfies a law of large numbers (see Proposition 2.5). Recall that a sequence R_n of probability measures on a topological space E satisfies a large deviation principle with rate function I whenever I is a nonnegative function having compact level sets and such that for every measurable A,

$$\inf_{x\in\operatorname{int}(A)} I(x) \le \liminf_{n\to\infty} \frac{1}{n}\log R_n(A) \le \limsup_{n\to\infty} \frac{1}{n}\log R_n(A) \le -\inf_{x\in\overline{A}} I(x), \tag{1.1}$$

where $\operatorname{int}(A)$ and $\overline{A}$ denote, respectively, the interior and the closure of A. For large deviations we refer to [8]. In this paper, we will show that a large deviation principle (LDP) holds for the sequence of laws of the measure-valued process X^n (abbreviated to X^n satisfies a LDP).

1.1. *The limitations of Poisson approximation.* The random allocations phenomenon is intimately connected with questions in Poisson approximation [1]. For any fixed t, X_t^n may be considered as the empirical measure Y_t^n of n identically distributed independent Poisson random variables with parameter t conditioned on the fact that their sum is equal to $\lfloor nt \rfloor$. By the Sanov theorem [8], Y_t^n satisfies a LDP in $\mathscr{P}(\mathbb{N})$ with rate function $H(\nu \mid \mathbf{p}_t) \stackrel{\Delta}{=} \sum_{i\in\mathbb{N}} \nu(i)\log\frac{\nu(i)}{\mathbf{p}_t(i)}$ [here $(\mathbf{p}_t(i))_{i\in\mathbb{N}}$ is the Poisson distribution with mean t]. As the probability that the sum of n independent Poisson random variables with parameter t equals $\lfloor nt \rfloor$ is of order $1/\sqrt{n}$, we immediately get the following LDP upper bound for X_t^n:

$$\limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(X_t^n \in C) \le -\inf_{\nu\in C} H(\nu \mid \mathbf{p}_t),$$

which by the way proves that, for fixed t, X_t^n satisfies a law of large numbers. A very natural question is whether the lower bound also holds. Analyzing the behavior of collections of dependent random variables by conditioning collections of independent random variables is a standard method in random combinatorics (see the characterization of the LDP for integer partitions profiles [7]). The ambitions of this paper consist first in providing with results of wide applicability and second in establishing a Sanov-like theorem in a dependent context.

As random allocations models arise in many applications ranging from algorithm analysis, random combinatorics and learning theory, they have been investigated by a variety of methods: mostly by combinatorial analysis (see [13, 12, 28, 5] and references therein) and probabilistic techniques using characteristic functions (see [18, 21–23] and references therein). Combinatorial analysis provides results of unrivaled precision, but tends to be involved when dealing with infinite-dimensional random variables. In classical probabilistic approaches, conditioning a collection of independent Poisson random variables and ingenious depoissonization arguments going back to [2] have often been used to obtain Central Limit Theorems.

This paper avoids the depoissonization approach, and characterizes the LDP for the flow of empirical occupancy measures by taking advantage of the simple dynamics of the model. We illustrate our main result by deriving a LDP for the degree distribution of sparse random graphs.

1.2. *Outline of the paper.* The main result of the paper, Theorem 2.7, is stated at the end of Section 2. The LDP upper bound is proved in Section 3 where a variational representation of the rate function is established. Thanks to Orlicz spaces techniques, the nonvariational representation of the rate function is established in Section 4. The LDP lower bound is established in Section 5 thanks to the classical change of measure argument. The difficulty lies in the construction of a rich enough collection of absolutely continuous changes of measures. In Section 6, the LDP for the flow of empirical occupancy measures is shown to hold with the same rate function when the topology is strengthened. In Section 7, a coupling argument allows deriving the LDP for the degree distribution of sparse random graphs from Theorem 2.7.

Convention. Henceforth, it is assumed that the value of all functions indexed by -1 are 0.

2. Main results.

2.1. *The model.* The dynamics of the process is as follows. If the kth ball is inserted into bin B^n_k which score $S^n_{k-1}(B^n_k) = i$, then

$$\begin{aligned} X^n_{k/n}(i+1) &= X^n_{(k-1)/n}(i+1) + 1/n, \\ X^n_{k/n}(i) &= X^n_{(k-1)/n}(i) - 1/n, \\ X^n_{k/n}(j) &= X^n_{(k-1)/n}(j), \qquad j \notin \{i, i+1\} \end{aligned}$$

and the value of the process remains constant on the time interval $[k/n, (k+1)/n)$. For any $i \geq 0$, each realization of the process $X^n(i)$ stands in the space $D([0,T], \mathbb{R})$ of right continuous left limited (cadlag) paths from $[0,T]$ to $\mathbb{R}$. The sample path space of X^n is $D_{\mathcal{P}} \triangleq D([0,T], \mathcal{P}(\mathbb{N}))$: the set of all ν: $[0,T] \mapsto \mathcal{P}(\mathbb{N})$ such that $\nu(i) \in D([0,T], \mathbb{R})$ for all $i \geq 0$.

Let us endow $D_{\mathcal{P}}$ with its canonical filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$ where $\mathcal{F}_t = \sigma(\pi_s; 0 \leq s \leq t)$ is generated by the canonical projections π_s: $\nu \in D_{\mathcal{P}} \mapsto \nu_s \in \mathcal{P}(\mathbb{N})$ and the σ-field on $\mathcal{P}(\mathbb{N})$ is induced by the usual product σ-field on $\mathbb{R}^{\mathbb{N}}$. Similarly, Ω^n is endowed with the natural filtration $(\mathcal{A}^n_t)_{0 \leq t \leq T}$ where $\mathcal{A}^n_t = \sigma(B^n_k; 1 \leq k \leq \lfloor nt \rfloor)$. The σ-fields on Ω^n and $D_{\mathcal{P}}$ are $\mathcal{A}^n_T$ and $\mathcal{F}_T$. Clearly, $X^n_t = X^n \circ \pi_t$, X^n is an $(\mathcal{A}^n_t)$-adapted process and the canonical process π is $(\mathcal{F}_t)$-adapted.

As the bins are chosen uniformly and independently at each time, the probability measure $\mathbb{P}^n$ on Ω^n is the product of the uniform distribution on $\{1, \dots, n\}$,

$$\mathbb{P}^n(d\omega) = \bigotimes_{1 \le k \le \lfloor nT \rfloor} \left(\frac{1}{n} \sum_{1 \le \alpha \le n} \delta_\alpha(d\omega_k) \right).$$

2.2. *The topologies.* Considering $\mathcal{P}(\mathbb{N})$ as a subset of sumable sequences $\ell_1(\mathbb{N})$, it is naturally endowed with the topology induced by ℓ_1 norm, that is, by the total variation metric

$$\|\pi - \rho\| = \sum_{i \ge 0} |\pi(i) - \rho(i)|, \qquad \pi, \rho \in \mathcal{P}(\mathbb{N}). \tag{2.1}$$

We shall not use the Skorokhod topology on the sample path space $D_{\mathcal{P}}$.

DEFINITION 2.1. The space $D_{\mathcal{P}}$ is endowed with the topology of uniform convergence associated with the norm

$$\|\nu - \mu\| \stackrel{\Delta}{=} \sup_{0 \le t \le T} \|\nu_t - \mu_t\|, \qquad \nu, \mu \in D_{\mathcal{P}}. \tag{2.2}$$

REMARK 2.2. Provided with this topology, $D_{\mathcal{P}}$ is a nonseparable complete metric space. The following result shows that we should not be distracted by separability, measurability and exponential tightness issues.

LEMMA 2.3. *There exists a compact subset $\mathcal{E}$ of $D_{\mathcal{P}}$ such that for all $n \ge 1$ and all $\omega \in \Omega^n$, $X^n(\omega)$ belongs to $\mathcal{E}$.*

PROOF. We take advantage of the simple form of the sample paths of $X^n(\omega)$. Let x^n be any realization of X^n. As for any t and n,

$$\sup_{t \le r, s < t+1/n} \|x^n_s - x^n_r\| = 2/n,$$

we have for any $0 \le t < t + \delta \le T$, $\sup_{t \le r, s < t+\delta} \|x^n_s - x^n_r\| \le (1 + n\delta)2/n = 2/n + 2\delta$.

On the other hand, for any t, the mean score per bin is $\sum_{i \ge 0} i x^n_t(i) \le t$. Hence, for all $t \le T$, x^n_t belongs to the relatively compact subset of $\mathcal{P}(\mathbb{N})$ consisting of the probability measures with their first moment bounded above by T. As $\mathcal{P}(\mathbb{N})$ is complete, this relatively compact subset is totally bounded.

As in the proof of Ascoli–Arzela's theorem, it follows from these considerations that for any $\varepsilon > 0$, one can build a finite collection of open balls of $D_{\mathcal{P}}$ with radius ε which covers $\bigcup_{n \ge 1} \{X^n(\omega); \omega \in \Omega^n\}$. This means that it is a totally bounded set in the complete metric space $D_{\mathcal{P}}$. Therefore, its closure $\mathcal{E}$ is compact. $\square$

REMARK 2.4. According to Lemma 4.1.5 in [8], if I is a rate function on $D_{\mathcal{P}}$ that is infinite outside $\mathcal{E}$, proving the LDP with rate function I in $\mathcal{E}$ endowed with the (complete, metrizable, separable) topology induced by the supremum norm is equivalent to proving the LDP in $D_{\mathcal{P}}$ provided with the topology induced by the same norm and the σ-algebra generated by canonical projections. This holds even if the σ-algebra defined on $D_{\mathcal{P}}$ is strictly smaller than the Borel σ-algebra.

2.3. *A larger class of models.* Suppose you observe a significant deviation of X^n from its limiting value (i.e., Poisson flow $\mathbf{p} = (\mathbf{p}_t)_{t\in[0,T]}$ according to the law of large numbers). One may ask for the most typical sample paths leading to this situation. As it will be shown that (X^n) satisfies a LDP, answering the previous question amounts to solving a variational problem whose objective function is the rate function of the LDP. The solutions of this variational problem are the limits in probability of the empirical occupancy measures arising in more general allocation schemes. In this section, we describe those general allocation schemes, and state the corresponding *laws of large numbers*. This will be useful when deriving the LDP lower bound (see Section 5). In this larger class of models, the choice of the bin B^n_{k+1} at time $k+1$ depends on the whole empirical distribution $X^n_{k/n}$ at time k. Let us take a continuous function λ on $[0, T] \times \mathbb{N}$ such that:

1. λ has range included in $[a, \infty)$ for some $a > 0$.
2. There exists some integer M such that $\lambda(t, i) = 1$ for all $i \geq M$ and all t.

Conditionally on $\mathcal{A}^n_{k/n}$, the probability of choosing a bin with score i is

$$\mathbb{Q}^n\big(S^n_k(B^n_{k+1}) = i \mid \mathcal{A}^n_{k/n}\big) = \lambda(k/n, i) X^n_{k/n}(i) / \langle \lambda_{k/n}, X^n_{k/n}\rangle, \tag{2.3}$$

where $\langle \lambda_{k/n}, X^n_{k/n}\rangle = \sum_{j\geq 0} \lambda(k/n, j) X^n_{k/n}(j)$. Let us remark that as $\inf \lambda \geq a > 0$, we have $\langle \lambda_{k/n}, X^n_{k/n}\rangle > 0$. The choice of the bin, among those of score i is uniform. Note that if $X^n_{k/n}(i; \omega) = 0$, then one cannot allocate the $(k+1)$th ball in a bin with score i. It is worth noting that under $\mathbb{Q}^n$, for any $d \geq M$, the $\mathbb{R}^d$-valued process formed by the projection of X^n on its first d coordinates is a Markov process. Indeed,

$$\langle \lambda_{k/n}, X^n_{k/n}\rangle = 1 + \sum_{i=0}^{M} \big(\lambda(k/n, i) - 1\big) X^n_{k/n}(i) \tag{2.4}$$

and hence for $d \geq M$, the law of the d-dimensional projection of X^n at time $(k+1)/n$ only depends on the value of the d-dimensional projection of X^n at time k/n. Under $\mathbb{Q}^n$, X^n is a projective limit of a vector-valued Markov process and satisfies the following law of large numbers whose proof is postponed to Section 5.

PROPOSITION 2.5 (Law of large numbers). *Let $\mathbb{Q}^n$ as before. Then, the sequence (X^n) converges, in probability, in $D_{\mathcal{P}}$ (see Definition 2.1) toward* *a deterministic process $\nu = (\nu_t)_{t\in[0,T]}$. This process is the unique solution of the following differential equation*:

$$\frac{d\nu_t}{dt}(i) = \ell_t^\nu(i-1)\nu_t(i-1) - \ell_t^\nu(i)\nu_t(i) \quad \textit{and} \quad \nu_0(i) = \delta_0(i), \quad i \in \mathbb{N}, \tag{2.5}$$

where $\ell_t^\nu(i) = \lambda(t,i)/\sum_{j\geq 0}\lambda(t,j)\nu_t(j)$.

In particular when $\mathbb{Q}^n = \mathbb{P}^n$ ($\ell^\nu \equiv 1$), the limiting path is the time-marginal flow of the Poisson process with parameter 1, denoted by $\mathbf{p}$, given by

$$\frac{d\mathbf{p}_t}{dt}(i) = \mathbf{p}_t(i-1) - \mathbf{p}_t(i). \tag{2.6}$$

2.4. *The rate function.* It will be convenient to associate with each path $\nu \in D_{\mathcal{P}}$ the relaxed measure on $[0,T]\times\mathbb{N}$,

$$\bar\nu(dt\,dz) = \nu_t(dz)\,dt.$$

A path $\nu \in D_{\mathcal{P}}$ is said to be *absolutely continuous* if for each $i \in \mathbb{N}$, there exists $\dot\nu(i)$ in $L_1([0,T],dt)$ such that $\nu_t(i) - \nu_0(i) = \int_{[0,t]}\dot\nu_s(i)\,ds$. For each absolutely continuous path ν, let us define υ^ν, $\bar\nu$-almost everywhere by

$$\upsilon_t^\nu(j) \triangleq -\sum_{i\leq j}\dot\nu_t(i) \quad \text{for } j \geq 0. \tag{2.7}$$

Let P be a probability measure and Q a nonnegative measure on some measure space. The *relative entropy* of Q with respect to P is defined by

$$H(Q|P) = \begin{cases} \mathbb{E}_Q \log\dfrac{dQ}{dP}, & \text{if } Q \text{ is a probability measure and } Q \ll P,\\ \infty, & \text{otherwise,}\end{cases} \tag{2.8}$$

with the convention $0\log 0 = 0$.

We are now in a position to define the rate function I, for any $\nu \in D_{\mathcal{P}}$,

$$I(\nu) \triangleq \begin{cases} \displaystyle\int_{[0,T]} H(\upsilon_t^\nu|\nu_t)dt, & \text{if } \nu \text{ is absolutely continuous},\\ \infty, & \text{otherwise.}\end{cases} \tag{2.9}$$

Note that $H(\cdot|P)$ is defined on the space of all nonnegative measures, but that if $H(Q|P) < \infty$ then Q is a probability measure (this conforms with the variational definition of relative entropy) (see [8], Lemma 6.2.13 and remarks thereafter). Hence for any ν satisfying $I(\nu) < \infty$, dt-almost everywhere υ_t^ν is a *probability* measure on $\mathbb{N}$. Moreover, using

$$\sum_{j\geq 0}\sum_{i>j}\nu_t(i) = \sum_{i\geq 0} i\,\nu_t(i) \quad \text{and} \quad \sum_{i>j}\dot\nu_t(i) = -\sum_{i\leq j}\dot\nu_t(i),$$

algebraic manipulations show that the time-derivative of the mean score is 1 dt-almost everywhere: $\sum_i i\,\dot\nu_t(i) = 1$. The finiteness of the rate function warrants that balls are allocated with unit intensity.

REMARK 2.6. The rate function representation may be interpreted in the following way: $v_t^\nu(j)$ represents the instantaneous rate of allocation into bins with score j at time t when following sample path [see (2.10)] ν, $H(v_t^\nu|\nu_t)$ represents the cost of following ν rather than satisfying the ordinary differential equation (2.6).

2.5. *The main results.* Let us recall that $D_{\mathcal{P}}$ is endowed with the topology defined by the uniform norm and the σ-algebra generated by canonical projections (see Remarks 2.2 and 2.4). The main result of the paper is the following theorem.

THEOREM 2.7. *The sequence $(X^n)_{n\geq 1}$ satisfies the LDP on $D_{\mathcal{P}}$ with the rate function I.*

As the identities (2.7) are equivalent to

$$\dot{\nu}_t(i) = v_t^\nu(i-1) - v_t^\nu(i) \quad \text{for } i \geq 0, \tag{2.10}$$

and as $I(\nu) = 0$ if and only if $v^\nu = \nu$, $\bar{\nu}$-almost everywhere, we obtain that $I(\nu) = 0$ if and only if for almost every t, $\dot{\nu}_t(i) = \nu_t(i-1) - \nu_t(i)$, for ν_t-almost all $i \geq 0$. Hence $\nu = \mathbf{p}$, which is in agreement with (2.6) in Proposition 2.5.

The rate function may actually be further interpreted.

PROPOSITION 2.8. *The function I is a convex rate function. Let $\nu \in D_{\mathcal{P}}$, then $I(\nu) < \infty$ if and only if ν is absolutely continuous and there exists a measurable $\mathbb{R}^{\mathbb{N}}$-valued function ℓ^ν which is defined $\bar{\nu}$-almost everywhere, such that*:

(i) *The following master equation is satisfied $\bar{\nu}$-almost everywhere*:

$$\dot{\nu}_t(i) = \ell_t^\nu(i-1)\nu_t(i-1) - \ell_t^\nu(i)\nu_t(i), \qquad i \geq 0. \tag{2.11}$$

(ii) $(\ell_t^\nu(i)\nu_t(i))_{i\geq 0}$ *defines a probability measure on $\mathbb{N}$, for dt-almost every t.*

(iii) $\int_{[0,T]}[\sum_{i=0}^\infty \nu_t(i)\ell_t^\nu(i)\log \ell_t^\nu(i)]\,dt < \infty$.

Alternative expressions for $I(\nu)$ are

$$\begin{aligned} I(\nu) &= \int_{[0,T]\times\mathbb{N}} \ell^\nu \log \ell^\nu \, d\bar{\nu} = \int_{[0,T]} \left[\sum_{i=0}^\infty \nu_t(i)\ell_t^\nu(i)\log \ell_t^\nu(i)\right] dt \\ &= \int_{[0,T]} H(\ell_t^\nu \nu_t \mid \nu_t)\, dt, \end{aligned}$$

where ℓ^ν is any process satisfying the above properties (i), (ii) *and* (iii).

PROOF. The convexity and the lower semicontinuity of I are direct consequences of the variational representation stated in Proposition 4.2.

If $I(\nu) < \infty$, then dt-almost everywhere υ_t^ν is a probability measure on $\mathbb{N}$ which is absolutely continuous with respect to ν_t. Let $\ell_t^\nu = \frac{d\upsilon_t^\nu}{d\nu_t}$ be its Radon–Nykodym derivative. Clearly, property (ii) holds. As

$$I(\nu) = \int_{[0,T]} H(\upsilon_t^\nu \mid \nu_t)\, dt = \int_{[0,T]} \langle \ell_t^\nu \log \ell_t^\nu, \nu_t \rangle\, dt, \tag{2.12}$$

property (iii) is satisfied. Finally, property (i) is given by (2.10).

Conversely, let ℓ_t^ν satisfy conditions (i), (ii) and (iii). Set $\upsilon_t^\nu = \ell_t^\nu \nu_t$. Then, (2.11) is (2.10) which is equivalent to (2.7). Finally, conditions (ii) and (iii) with (2.12) imply that $I(\nu)$ is finite. $\square$

Let ν be an absolutely continuous path. If $\nu_t(i) > 0$, (2.11) gives $\ell_t^\nu(i) = [-\sum_{j \le i} \dot{\nu}_t(j)]/\nu_t(i)$, so that $\ell_t^\nu(i)$ is uniquely defined up to $\bar{\nu}$-a.e. equality on $\{(t,i); \nu_t(i) > 0\}$. On the other hand, (2.11) and (2.12) are insensitive to the values of ℓ^ν on the complementary set $\{(t,i); \nu_t(i) = 0\}$. Therefore,

$$\ell_t^\nu(i) = \begin{cases} \left[-\sum_{j \le i} \dot{\nu}_t(j)\right]/\nu_t(i), & \text{if } \nu_t(i) > 0, \\ 1, & \text{if } \nu_t(i) = 0, \end{cases} \tag{2.13}$$

is a useful measurable inversion formula for ℓ^ν.

3. The upper bound.

3.1. *Statement of the upper bound.* As we will resort to duality arguments in Sections 4 and 5, let us first define an ad hoc set of test functions. This set of test functions will be rich enough to allow us to use the variational characterization of entropy when we will identify the rate function. Moreover, these functions will be sufficiently regular in order to simplify the proofs. Let g be a real function on $\mathbb{N}$ (a sequence of real numbers); we set $Dg(j) \stackrel{\Delta}{=} g(j+1) - g(j)$, for all $j \ge 0$. For any function $G\colon [0,T] \times \mathbb{N} \to \mathbb{R}$, let us denote for all $j \in \mathbb{N}$, $G(j)\colon t \in [0,T] \mapsto G_t(j) \stackrel{\Delta}{=} G(t,j)$ and for all $0 \le t \le T$, $G_t = (G_t(j))_{j \ge 0}$. The set of relevant test functions is

$$\mathcal{G} \stackrel{\Delta}{=} \Big\{G\colon [0,T] \times \mathbb{N} \to \mathbb{R}; \sup_{t,j} |DG_t(j)| < \infty, G(j) \in \mathcal{C}, \forall\, j \in \mathbb{N}\Big\},$$

where $\mathcal{C}$ is the space of all functions $f\colon [0,T] \mapsto \mathbb{R}$ which are absolutely continuous and such that $f(T) = 0$. For any G in $\mathcal{G}$, we will denote by $\dot{G}_t$ the generalized derivative of G_t with respect to t; that is,

$$G_t(j) = -\int_{[t,T]} \dot{G}_s(j)\, ds, \qquad t \in [0,T],\ j \in \mathbb{N}.$$

Let us also introduce the notation $\dot{\nu}(G)$. For all $G \in \mathcal{G}$ and $\nu \in D_{\mathcal{P}}$,

$$\dot{\nu}(G) \stackrel{\Delta}{=} -\langle G_0, \nu_0 \rangle - \int_{[0,T]} \langle \dot{G}_t, \nu_t \rangle\, dt. \tag{3.1}$$

The main result of the section is the following variational formulation of the large deviation upper bound.

PROPOSITION 3.1. *Let $D_{\mathcal{P}}$ be endowed with the topology defined by the uniform norm (see Definition 2.1). For any closed measurable subset C of $D_{\mathcal{P}}$ we have*

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}^n(X^n\in C)\le-\inf_{\nu\in C}\sup_{G\in\mathcal{G}}\left\{\dot{\nu}(G)-\int_{[0,T]}\log\langle\exp(DG_t),\nu_t\rangle\,dt\right\}.$$

PROOF. In Lemma 3.3 below, the upper bound is proved for C measurable convex and compact. The convexity restriction is removed using MinMax argument inspired from [27], Theorem 4.1. As by Lemma 2.3, the laws of the X^n's are compactly supported, the upper bound holds for all measurable closed subsets of $D_{\mathcal{P}}$ or $\mathcal{E}$. $\square$

3.2. *Exponential martingale.* We introduce a family of exponential martingales $Z^{G,n}$ which will allow us, by means of Lemma 3.2 below, to derive in Lemma 3.3 the upper bound for compact convex subsets.

For any G in $\mathcal{G}$ and $n\ge 1$, let us define the process $Z^{G,n}$ by

$$\begin{aligned}\frac{1}{n}\log Z_t^{G,n}&\triangleq\langle G_t,X_t^n\rangle-\langle G_0,X_0^n\rangle-\int_{[0,t]}\langle\dot{G}_s,X_s^n\rangle\,ds\\&\quad-\sum_{k=0}^{\lfloor nt\rfloor-1}\frac{1}{n}\log\sum_{j\ge0}\exp\bigl(DG_{(k+1)/n}(j)\bigr)X_{k/n}^n(j).\end{aligned}$$

As DG is bounded, there exists $c\ge 0$ such that $|G(i)|\le c(1+i)$ for all $i\in\mathbb{N}$. As for all $0\le t\le T$, $\sum_{i\ge0}iX_t^n(i)\le T$, all the terms in the definition of $Z^{G,n}$ are well defined and $Z^{G,n}$ is a bounded process.

LEMMA 3.2. *For any $G\in\mathcal{G}$ and $n\ge1$, $(Z_t^{G,n})_{0\le t\le T}$ is a $\mathbb{P}^n$-martingale with respect to the filtration $(\mathcal{A}_t^n)_{0\le t\le T}$. In particular, $\mathbb{E}_{\mathbb{P}^n}Z_T^{G,n}=1$.*

PROOF. It is enough to check that for any $0\le t\le t+h\le T$,

$$\mathbb{E}_{\mathbb{P}^n}(Z_{t+h}^{G,n}/Z_t^{G,n}\mid\mathcal{A}_t^n)=1.$$

We have

$$\begin{aligned}\tfrac{1}{n}\log[Z_{t+h}^{G,n}/Z_t^{G,n}]&=\langle G_{t+h},X_{t+h}^n\rangle-\langle G_t,X_t^n\rangle-\int_{[t,t+h]}\langle\dot{G}_s,X_s^n\rangle\,ds\\&\quad-\sum_{k=\lfloor nt\rfloor}^{\lfloor n(t+h)\rfloor-1}\frac{1}{n}\log\sum_{j\ge0}\exp\bigl(DG_{(k+1)/n}(j)\bigr)X_{k/n}^n(j).\end{aligned}\tag{3.2}$$

Note that if $\frac{k}{n} \le t < \frac{k+1}{n} \le t+h$, then

$$\begin{aligned}\langle G_{t+h}, X^n_{t+h}\rangle - \langle G_t, X^n_t\rangle \\ &= \int_{[t,t+h]} \langle \dot G_s, X^n_s\rangle\, ds + \langle G_{(k+1)/n}, X^n_{(k+1)/n}\rangle - \langle G_{(k+1)/n}, X^n_{k/n}\rangle.\end{aligned}$$

If $\lfloor n(t+h)\rfloor = \lfloor nt\rfloor$, the right-hand side vanishes and there is nothing to prove. Using cascade conditioning, all other cases reduce to $\lfloor n(t+h)\rfloor = \lfloor nt\rfloor + 1$. Furthermore, it is enough to consider the case $\lfloor nt\rfloor = nt$ and $1/n \le h < 2/n$.

Hence the right-hand side of (3.2) reduces to

$$\langle G_{(k+1)/n}, X^n_{(k+1)/n}\rangle - \langle G_{(k+1)/n}, X^n_{k/n}\rangle - \frac{1}{n}\log \sum_{j\ge 0} \exp\bigl(DG_{(k+1)/n}(j)\bigr) X^n_{k/n}(j).$$

The proof is completed by noticing

$$\begin{aligned}&\mathbb{E}_{\mathbb{P}^n}\bigl[\exp(n[\langle G_{(k+1)/n}, X^n_{(k+1)/n}\rangle - \langle G_{(k+1)/n}, X^n_{k/n}\rangle])|\mathcal{A}^n_t\bigr] \\ &\quad = \sum_{j\ge 0} \exp\bigl(DG_{(k+1)/n}(j)\bigr) X^n_{k/n}(j). \qquad \square\end{aligned}$$

3.3. *Compact convex subsets.* The LDP upper bound for convex compact sets is established thanks to a general min–max theorem due to Sion [26] (see also [8], Exercises 2.2.38, 4.5.5 for applications of this theorem to the derivation of LDP upper bounds). Let $C \subset D_{\mathcal{P}}$ be a convex compact set, for $G \in \mathcal{G}$ and $\nu \in D_{\mathcal{P}}$ define the two functionals,

$$\mathcal{K}(G,\nu) \triangleq \dot\nu(G) - \int_{[0,T]} \log\langle \exp(DG_t), \nu_t\rangle\, dt,$$

$$\mathcal{K}^n(G,\nu) \triangleq \dot\nu(G) - \sum_{k=0}^{\lfloor nT\rfloor -1} \frac{1}{n} \log\langle \exp(DG_{(k+1)/n}), \nu_{k/n}\rangle.$$

The two functionals are convex with respect to ν thanks to the concavity of log, and concave with respect to G thanks to Hölder's inequality and the fact that log is increasing. The continuity with respect to ν and the upper semicontinuity with respect to G follow from the definition of $\mathcal{G}$.

LEMMA 3.3. *Let C be a measurable convex compact subset of $D_{\mathcal{P}}$, then*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}^n(X^n \in C) \le -\inf_{\nu\in C} \sup_{G\in\mathcal{G}} \mathcal{K}(G,\nu).$$

PROOF. By an exponential Markov inequality, for any $G \in \mathcal{G}$,

$$\mathbb{P}^n(X^n \in C) \le \mathbb{P}^n\bigl(e^{n\mathcal{K}^n(G,X^n)} \ge \inf_{\nu\in C} e^{n\mathcal{K}^n(G,\nu)}\bigr) \le \exp\bigl(-n \inf_{\nu\in C} \mathcal{K}^n(G,\nu)\bigr),$$

since as $\mathbb{E}_{\mathbb{P}^n} e^{n\mathcal{K}^n(G,X^n)} = \mathbb{E}_{\mathbb{P}^n} Z_T^{G,n}$, by Lemma 3.2 we have $\mathbb{E}_{\mathbb{P}^n} e^{n\mathcal{K}^n(G,X^n)} = 1$. We may now optimize with respect to $G \in \mathcal{G}$:

$$\mathbb{P}^n(X^n \in C) \leq \inf_G \exp\Big(-n \inf_{\nu\in C} \mathcal{K}^n(G,\nu)\Big) = \exp\Big(-n \sup_{G\in\mathcal{G}} \inf_{\nu\in C} \mathcal{K}^n(G,\nu)\Big).$$

Letting n tend to infinity, one obtains

$$\limsup_n \frac{1}{n} \log \mathbb{P}^n\left(X^n \in C\right) \leq -\liminf_n \sup_G \inf_{\nu\in C} \mathcal{K}^n(G,\nu).$$

Let us prove that

$$\liminf_n \sup_{G\in\mathcal{G}} \inf_{\nu\in C} \mathcal{K}^n(G,\nu) \geq \sup_G \inf_{\nu\in C} \mathcal{K}(G,\nu). \tag{3.3}$$

As DG is bounded and t-continuous and ν is right continuous, $\mathcal{K}^n$ converges pointwise toward $\mathcal{K}$. For any fixed $G \in \mathcal{G}$, the sequence $\mathcal{K}^n(G,\cdot)$ converges uniformly toward $\mathcal{K}(G,\cdot)$ on the compact set C. Let us check that for a fixed G, for every $\varepsilon > 0$, for n sufficiently large, for all ν in C,

$$\left|\int_{[0,T]} \left[\log\langle \exp(DG_{(\lfloor nt\rfloor+1)/n}), \nu_{\lfloor nt\rfloor/n}\rangle - \log\langle \exp(DG_t), \nu_t\rangle\right] dt\right| \leq \varepsilon.$$

Indeed, as we may assume that there exists some L such that $\exp(-L) \leq \exp(DG_t(i)) \leq \exp(L)$ for all $t \in [0,T]$ and i,

$$\begin{aligned}&\left|\int_{[0,T]} \left[\log\langle \exp(DG_{(\lfloor nt\rfloor+1)/n}), \nu_{\lfloor nt\rfloor/n}\rangle - \log\langle \exp(DG_t), \nu_t\rangle\right] dt\right| \\ &\qquad \leq \int_{[0,T]} \left[e^{2L}\|\nu_{\lfloor nt\rfloor/n} - \nu_t\| + \langle |\exp(DG_{(\lfloor nt\rfloor+1)/n}) - \exp(DG_t)|, \nu_t\rangle\right] dt.\end{aligned}$$

The first term in the integrand tends uniformly towards 0 on the compact set C since compact sets are equicontinuous. The second term in the integrand tends uniformly toward 0 because ν_t are uniformly tight thanks to the compact containment property of compacta of $D_{\mathcal{P}}$ and because each $DG_t(i)$ is absolutely continuous with respect to t.

Let us take $\varepsilon > 0$. For any n and G, let $\nu^{G,n} \in C$ be such that $\mathcal{K}^n(G,\nu^{G,n}) \leq \inf_{\nu\in C} \mathcal{K}^n(G,\nu) + \varepsilon$. Because of uniform convergence, for any G, there exists $n^G \geq 1$ such that for all $n \geq n^G$: $\inf_{\nu\in C} \mathcal{K}^n(G,\nu) \geq \mathcal{K}^n(G,\nu^{G,n}) - \varepsilon \geq \mathcal{K}(G,\nu^{G,n}) - 2\varepsilon \geq \inf_{\nu\in C} \mathcal{K}(G,\nu) - 2\varepsilon$. Hence,

$$\sup_{G\in\mathcal{G}} \liminf_{n\to\infty} \inf_{\nu\in C} \mathcal{K}^n(G,\nu) \geq \sup_{G\in\mathcal{G}} \inf_{\nu\in C} \mathcal{K}(G,\nu).$$

As, $\liminf_{n\to\infty} \sup_{G\in\mathcal{G}} \geq \sup_{G\in\mathcal{G}} \liminf_{n\to\infty}$, this proves (3.3).

Applying Sion's Theorem [26], the right-hand side in (3.3) is identified with $\inf_{\nu\in C} \sup_{G\in\mathcal{G}} \mathcal{K}(G,\nu)$. $\square$

4. The rate function. Proposition 4.2 below identifies the rate function appearing in Proposition 3.1 with the rate function I defined at (2.9). It will be proved using the Riesz representation theorem in Orlicz spaces. Using the Riesz representation theorem in L_2 would have been appropriate if we were facing a Gaussian situation, but the bins and balls model resorts to Poisson approximation. Orlicz spaces constitute a tailor-made framework in such a case [14]. For the sake of completeness, let us first recall some basic facts about Orlicz.

4.1. *Orlicz spaces.* A *Young function* θ is an even, convex, $[0,\infty]$-valued function satisfying $\theta(0)=0$, $\lim_{s\to+\infty}\theta(s)=+\infty$ and $\theta(s_0)<+\infty$ for some $s_0>0$. The convex conjugate θ^* of the Young function θ: $\theta^*(t)=\sup_{s\in\mathbb{R}}\{st-\theta(s)\}$ is also a Young function, and the Young inequality states $st\le\Lambda(s)+\Lambda^*(t)$. In the sequel, the relevant Young functions are $\tau(x)\stackrel{\Delta}{=}\exp(|x|)-|x|-1$ and $\tau^*(x)=(|x|+1)\log(|x|+1)-|x|$.

Let μ be a nonnegative bounded measure on the measurable space $(\Sigma,\mathcal{A})$. Consider the following vector spaces, where μ-almost everywhere equal functions are identified:

$$L_\theta=\left\{f\colon\ \Sigma\to\mathbb{R},\exists\, a>0,\int_\Sigma\theta\left(\frac{f}{a}\right)d\mu<\infty\right\},$$
$$M_\theta=\left\{f\colon\ \Sigma\to\mathbb{R},\forall\, a>0,\int_\Sigma\theta\left(\frac{f}{a}\right)d\mu<\infty\right\}.$$

The Orlicz space associated with θ is the Banach space induced by the following norm on L_θ (see [24] and references therein):

$$\|f\|_\theta=\inf\left\{a>0,\int_\Sigma\theta\left(\frac{f}{a}\right)d\mu\le 1\right\}. \tag{4.1}$$

Hölder's inequality holds between L_θ and the Orlicz space L_{θ^*},

$$\forall\, f\in L_\theta, g\in L_{\theta^*},\qquad fg\in L^1(\mu)\quad\text{and}\quad\int_\Sigma|fg|\,d\mu\le 2\|f\|_\theta\|g\|_{\theta^*}. \tag{4.2}$$

Thus any g in L_{θ^*} defines a continuous linear form on L_θ for the duality bracket $\langle f,g\rangle=\int fg\,d\mu$. Although, the topological dual space of $(L_\theta,\|\cdot\|_\theta)$ may be larger than L_{θ^*}, we always have the following version of the Riesz representation theorem. (See, e.g., [15], Section 4, for a proof.)

THEOREM 4.1. *Let θ be a finite Young function and θ^* its convex conjugate. The topological dual space of M_θ can be identified, using the previous duality bracket, with L_{θ^*}: $M'_\theta\simeq L_{\theta^*}$.*

4.2. *Variational representation of the rate function.*

PROPOSITION 4.2. *For every $\nu \in D_{\mathcal{P}}$, we have*

$$I(\nu) = \sup_{G \in \mathcal{G}} \left\{ \dot{\nu}(G) - \int_{[0,T]} \log\langle \exp(DG_t), \nu_t \rangle \, dt \right\}.$$

PROOF. Let ν belong to $D_{\mathcal{P}}$. By (4.1), for any $G \in \mathcal{G}$,

$$\|DG\|_{\tau,\bar{\nu}} = \inf\left\{ a > 0; \int_{[0,T]\times\mathbb{N}} \tau(DG/a)\, d\bar{\nu} \le 1 \right\}.$$

Let $K(\nu) \stackrel{\Delta}{=} \sup_{G\in\mathcal{G}} \mathcal{K}(G,\nu) = \sup_{G\in\mathcal{G}}\{\dot{\nu}(G) - \int_{[0,T]} \log\langle \exp(DG_t), \nu_t\rangle\, dt\}$, so that for any $a > 0$ and $G \in \mathcal{G}$: $\dot{\nu}(G/a) \le K(\nu) + \int_{[0,T]} \log\langle \exp(DG_t/a), \nu_t\rangle\, dt$. Subtracting $\int_{[0,T]} \langle DG_t/a, \nu_t\rangle\, dt$ from both sides,

$$\begin{aligned}
\dot{\nu}(G/a) &- \int_{[0,T]} \langle DG_t/a, \nu_t\rangle\, dt \\
&\le K(\nu) + \int_{[0,T]} \big[\log\langle \exp(DG_t/a), \nu_t\rangle - \langle DG_t/a, \nu_t\rangle\big]\, dt \\
&\stackrel{\text{(a)}}{\le} K(\nu) + \int_{[0,T]} \langle \exp(DG_t/a) - DG_t/a - 1, \nu_t\rangle\, dt \\
&\stackrel{\text{(b)}}{\le} K(\nu) + \int_{[0,T]} \langle \tau(DG_t/a), \nu_t\rangle\, dt,
\end{aligned}$$

where (a) comes from $\log x \le x - 1$ and (b) from $\exp(x) - x - 1 \le \tau(x)$. Choosing $a = \|DG\|_{\tau,\bar{\nu}}$,

$$\dot{\nu}(G) - \int_{[0,T]\times\mathbb{N}} DG\, d\bar{\nu} \le [K(\nu) + 1]\|DG\|_{\tau,\bar{\nu}}.$$

As an analogue inequality can be proved by replacing G by $-G$, and as by Hölder's inequality (4.2), $|\int_{[0,T]\times\mathbb{N}} DG\, d\bar{\nu}| \le 2\|\mathbb{1}_{[0,T]}\|_{\tau^*,\bar{\nu}}\|DG\|_{\tau,\bar{\nu}} = 2T\tau^*(1)\|DG\|_{\tau,\bar{\nu}}$,

$$|\dot{\nu}(G)| \le [K(\nu) + 1 + 2T\tau^*(1)]\|DG\|_{\tau,\bar{\nu}}.$$

Let us first prove that if $K(\nu) < \infty$, then $K(\nu) = I(\nu)$. Now, assume ν is such that $K(\nu) < \infty$. The above estimate implies first that for all F, G in $\mathcal{G}$, if $DF = DG$ then $\dot{\nu}(F) = \dot{\nu}(G)$ and second that $\dot{\nu}$ is a continuous linear form on $M_\tau(\bar{\nu})$. By Theorem 4.1, there exists ℓ^ν in $L_{\tau^*}(\bar{\nu})$ such that

$$\dot{\nu}(G) = \int_{[0,T]} \langle \ell_t^\nu DG_t, \nu_t\rangle\, dt \qquad \forall\, G \in \mathcal{G}. \tag{4.3}$$

It follows that

$$\begin{aligned}
K(\nu) &= \sup_{G\in\mathcal{G}}\left\{\dot\nu(G) - \int_{[0,T]} \log\langle \exp(DG_t), \nu_t\rangle\, dt\right\}\\
&\overset{(a)}{=} \sup_{G\in\mathcal{G}}\left\{\int_{[0,T]} \big(\langle \ell_t^\nu DG_t, \nu_t\rangle - \log\langle \exp(DG_t), \nu_t\rangle\big)\, dt\right\}\\
&\overset{(b)}{=} \sup_{g\in\ell^\infty}\left\{\int_{[0,T]} (\langle g, \ell_t^\nu \nu_t\rangle - \log\langle e^g, \nu_t\rangle)\, dt\right\}\\
&\overset{(c)}{=} \int_{[0,T]} \sup_{g\in\ell^\infty}\{\langle g, \ell_t^\nu \nu_t\rangle - \log\langle e^g, \nu_t\rangle\}\, dt\\
&\overset{(d)}{=} \int_{[0,T]} H(\ell_t^\nu \nu_t \mid \nu_t)\, dt.
\end{aligned}$$

Equation (a) follows directly from (4.3) whereas equation (b) follows by setting $g = DG_t \in \ell^\infty$. Equation (c) follows from Theorem 2 in [25]. This theorem states that, under mild assumptions, in the conjugate computation of a convex integral functional one may exchange integral and supremum. Observe that this theorem is obvious for a measure supported by a finite set. Finally, equation (d) follows from the variational representation of the relative entropy ([8], Lemma 6.2.13). Note that if $K(\nu) < \infty$, then equation (c) implies that dt-almost everywhere $\ell_t^\nu \nu_t$ defines a probability measure on $\mathbb{N}$.

By Proposition 2.8, completing the proof of $K(\nu) = I(\nu)$ when $K(\nu) < \infty$ amounts to checking that the master equation is (2.11). Choosing $G_t^{\varphi,i}(j) = \varphi_t \mathbb{1}_{\{i\}}(j)$ where φ is continuously differentiable and $\varphi_T = 0$, with (3.1),

$$\dot\nu(G^{\varphi,i}) = -\varphi_0 \nu_0(i) - \int_{[0,T]} \dot\varphi_t \nu_t(i)\, dt$$

and (4.3) leads us to

$$\dot\nu(G^{\varphi,i}) = \int_{[0,T]} \varphi_t[\ell_t^\nu(i-1)\nu_t(i-1) - \ell_t^\nu(i)\nu_t(i)]\, dt.$$

As these identities hold for all φ and i, (2.11) is satisfied and thus $K(\nu) = I(\nu)$ whenever $K(\nu) < \infty$.

Finally, if $I(\nu) < \infty$, by Proposition 2.8 we have (2.11) which is equivalent to (4.3) by the above-described computation. Now, according to the computation following (4.3), we obtain that $K(\nu) = I(\nu)$. $\square$

A simple corollary of Proposition 2.8 is the following.

COROLLARY 4.3. *If ν satisfies $I(\nu) < \infty$, then for all $0 \le s \le t \le T$,*

$$\|\nu_t - \nu_s\| \le 2(t-s).$$

As a matter of fact, the effective domain of the rate function I is included in the compact subset of $D_{\mathcal{P}}$ mentioned in Lemma 2.3.

5. The lower bound.

In this section we prove the following lower bound.

PROPOSITION 5.1. *For any open measurable subset U of $D_{\mathcal{P}}$ we have*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}^n (X^n \in U) \geq - \inf_{\nu \in U} I(\nu).$$

Without loss of generality, $I(U) \stackrel{\Delta}{=} \inf_{\nu\in U} I(\nu)$ is assumed to be finite. The nth change of measure associated with path $\nu \in U$ satisfying $I(\nu) < \infty$ is the twisted probability $\mathbb{Q}^{\nu,n}$ defined by (2.3) with $\lambda = \ell^\nu$, provided that ℓ^ν is regular enough. Henceforth, P^n and $Q^{\nu,n}$ stand for the laws of X^n under $\mathbb{P}^n$ and $\mathbb{Q}^{\nu,n}$ (see Section 2). The changes of measure are used through the following device. For any $\varepsilon > 0$,

$$\begin{aligned} \frac{1}{n} \log \mathbb{P}^n \left(X^n \in U\right) &= \frac{1}{n} \log \mathbb{E}_{\mathbb{Q}^{\nu,n}} \left(\frac{d\mathbb{P}^n}{d\mathbb{Q}^{\nu,n}} \mathbb{1}_{\{X^n \in U\}} \right) \\ &\geq \inf_{\xi\in V} \frac{1}{n} \log \frac{dP^n}{dQ^{\nu,n}}(\xi) + \frac{1}{n} \log \mathbb{Q}^{\nu,n}(X^n \in V) \\ &\geq \frac{1}{n} \log \frac{dP^n}{dQ^{\nu,n}}(\nu) - \varepsilon - \varepsilon \end{aligned} \tag{5.1}$$

for any small enough neighborhood V of ν with $V \subset U$ and any large enough n, provided that $\mathbb{Q}^{\nu,n}$ satisfies the three following properties:

(α) For any open neighborhood V of ν, $\lim_n \frac{1}{n} \log \mathbb{Q}^{\nu,n}(X^n \in V) = 0$.
(β) P^n is absolutely continuous with respect to $Q^{\nu,n}$.
(γ) The map $\xi \in V \mapsto \frac{dP^n}{dQ^{\nu,n}}(\xi)$ is continuous at ν.

Membership in $\mathcal{D}_I$, the effective domain of I, does not warrant these three properties. Therefore we focus on a subset of $\mathcal{D}_I$: the set of nice ν's.

5.1. *The nice ν's.*

Property (α) will follow from the law of large numbers (Proposition 2.5). Property (β) will be easily checked if bounds are imposed on ν. Property (γ) will be checked if ν and ℓ^ν are sufficiently smooth. We will assume that ν is nice according to the following definition.

DEFINITION 5.2. A path $\nu \in D_{\mathcal{P}}$ is said to be *nice* if:

(i) ν belongs to $\mathcal{D}_I$; that is, $I(\nu) < \infty$.
(ii) For all $t > 0$ and $i \in \mathbb{N}$, $\nu_t(i) > 0$.
(iii) There exists $M \geq 0$ such that for all $i \geq M$ and all $0 \leq t \leq T$, $\ell_t^\nu(i) = 1$ and there exists $\beta > 0$ such that for all $t > 0$ and $i \in \mathbb{N}$, $\ell_t^\nu(i) \geq \beta$.
(iv) For all $i \in \mathbb{N}$, $\ell_t^\nu(i)$ is C^2 with respect to t.

Under conditions (ii) and (iv), formula (2.13) allows determining $\ell_t^\nu(i)$ as a function of ν for all i and $t > 0$.

Let ν be nice; $\mathbb{Q}^{\nu,n}$ is the nth *twisted probability measure* associated with it; it is defined by (2.3) with $\lambda = \ell^\nu$.

The main property of ν and $\mathbb{Q}^{\nu,n}$ when ν is nice is stated in the following lemma.

LEMMA 5.3 (Nice ν's are really nice). *For any nice* ν,

$$\sup_V \liminf_n Q^{\nu,n} \operatorname*{ess\,inf}_{\xi\in V}\left(\frac{1}{n}\log\frac{dP^n}{dQ^{\nu,n}}(\xi)\right) \geq -I(\nu),$$

where the supremum is taken over all open measurable neighborhoods V of ν.

This lemma together with (5.1) leads to the desired lower bound for any open measurable neighborhood of any nice ν. In order to extend this result to the general case, the following density result will be needed.

LEMMA 5.4 (Nice ν's are dense). *For each $\nu \in \mathcal{D}_I$, there exists a sequence $(\nu_m)_{m\geq 1}$ of nice sample paths such that* $\lim_{m\to\infty}\sup_{0\leq t\leq T}\|\nu_t - (\nu_m)_t\| = 0$ *and* $\lim_{m\to\infty} I(\nu_m) = I(\nu)$.

The proofs of these lemmas are postponed to after the proof of Proposition 5.1.

5.2. *Proof of the lower bound.* Assuming Proposition 2.5, Lemma 5.3 and Lemma 5.4, we can give a proof of lower bound.

PROOF OF PROPOSITION 5.1. Let U be any open measurable subset of $D_{\mathcal{P}}$ with $I(U) < \infty$. For any $\varepsilon > 0$, let $\nu_* \in U$ be such that $I(\nu_*) < I(U) + \varepsilon$. By Lemma 5.4, there exists a sequence of nice sample paths $(\nu_m)_{m\geq 1}$ converging to ν_* in U such that $I(\nu_m)$ converges towards $I(\nu_*)$. Hence, there exists a nice ν in U such that $I(\nu) < I(\nu_*) + \varepsilon < I(U) + 2\varepsilon$. Let $Q^{\nu,n}$ be the nth twisted probability law of X^n associated with ν, then

$$\begin{aligned}
&\liminf_n \frac{1}{n}\log \mathbb{P}^n(X^n \in U)\\
&\quad\overset{(a)}{\geq} \sup_{V:\,\nu\in V\subset U}\biggl[\liminf_n Q^{\nu,n}\text{-}\operatorname*{ess\,inf}_{\xi\in V}\left(\frac{1}{n}\log\frac{dP^n}{dQ^{\nu,n}}(\xi)\right)\\
&\qquad\qquad\qquad\qquad + \liminf_n \frac{1}{n}\log Q^{\nu,n}(V)\biggr]\\
&\quad\overset{(b)}{\geq} \sup_{V:\,\nu\in V\subset U} \liminf_n Q^{\nu,n}\text{-}\operatorname*{ess\,inf}_{\xi\in V}\left(\frac{1}{n}\log\frac{dP^n}{dQ^{\nu,n}}(\xi)\right)\\
&\quad\overset{(c)}{\geq} -I(\nu) \geq -I(U) - 2\varepsilon,
\end{aligned}$$

where (a) follows from (5.1), (b) from Proposition 2.5 and (c) from Lemma 5.3. □

5.3. *Proof of Proposition 2.5.* Recall that for all $n \geq 1$, under $\mathbb{Q}^{\nu,n}$, X^n is still a Markov process (as pointed out in Section 2.3). The argument relies on diffusion approximation techniques. As limiting distributions are degenerate, it is enough to rely on the following version of Corollary 4.2 in [9], page 355. In the sequel, if A is a matrix, $A^\dagger$ denotes its transpose.

LEMMA 5.5. *Let Y^n be a sequence of $\mathbb{R}^d$-valued Markov chains with initial condition distributed according to μ and vanishing maximal jump size. If the first-order differential equation $\frac{dy}{dt} = b(y,t)$ has a unique solution in $C^1([0,T], \mathbb{R}^d)$ for any initial condition y_0 in the support of μ, then if the sequences*

$$b_n(y,t) \stackrel{\Delta}{=} n\mathbb{E}[Y^n_{k+1} - Y^n_k \mid Y^n_k = y, k = \lfloor nt \rfloor],$$

$$a_n(y,t) \stackrel{\Delta}{=} n\mathbb{E}[(Y^n_{k+1} - Y^n_k)^\dagger (Y^n_{k+1} - Y^n_k) \mid Y^n_k = y, k = \lfloor nt \rfloor],$$

satisfy for each r and T,

$$\text{(i)} \quad \lim_n \sup_{\|y\| \leq r, t \leq T} |a_n(y,t)| = 0,$$

$$\text{(ii)} \quad \lim_n \sup_{\|y\| \leq r, t \leq T} |b_n(y,t) - b(y,t)| = 0,$$

then the sequence of processes $t \mapsto Y^n_{\lfloor nt \rfloor}$ converges in law towards the process Y with Y_0 distributed according to μ and for each y_0 in the support set of μ, if $Y_0 = y_0$, Y is the unique solution of $\frac{dy}{dt} = b(y,t)$ with initial condition y_0.

PROOF OF PROPOSITION 2.5. By Lemma 2.3, the sequence $(Q^{\nu,n})_{n \geq 1}$ is tight in $D_{\mathcal{P}}$. Using observation (2.4), it is thus enough to check the conditions of Lemma 5.5 for the d-dimensional projections of X^n for all $d > M$.

The limiting ordinary differential equation is:

$$\frac{d\xi_t(i)}{dt} = \frac{1}{\langle \ell_t \xi_t \rangle} [\ell^\nu_t(i-1)\xi_t(i-1) - \ell^\nu_t(i)\xi_t(i)], \qquad i \in \mathbb{N}.$$

As ν is nice, $\ell^\nu_t(i)$ is bounded from below by some $\beta > 0$. Hence the limiting differential equation satisfies the local Lipschitz condition and has a unique solution.

Now we have to check conditions (i) and (ii) in Lemma 5.5. For all $i \in \mathbb{N}$: the ith coordinate of $b_n(x,t)$ equals

$$n\mathbb{E}[X^n_{t+1/n}(i) - X^n_t(i) \mid X^n_t = x] = \frac{1}{\langle \ell_t, x \rangle} [\ell^\nu_t(i-1)\, x(i-1) - \ell^\nu_t(i) x(i)],$$

where expectation is taken with respect to $\mathbb{Q}^{\nu,n}$. Hence condition (ii) is enforced.

The conditional covariance matrix is symmetric tridiagonal. The diagonal and off-diagonal terms satisfy

$$n\mathbb{E}\big[\mathrm{Cov}(X^n_{t+1/n} - X^n_t)[i,i] \mid X^n_t = x\big] = \frac{\ell^\nu_t(i)x(i) + \ell^\nu_t(i-1)x(i-1)}{n\langle \ell^\nu_t, x\rangle},$$

$$n\mathbb{E}\big[\mathrm{Cov}(X^n_{t+1/n} - X^n_t)[i,i+1] \mid X^n_t = x\big] = -\frac{\ell^\nu_t(i)x(i)}{n\langle \ell^\nu_t, x\rangle}.$$

The sum of the absolute values of the coefficients of $a_n(x,t)$ is bounded by $4/n$, which warrants condition (i). $\square$

5.4. *Proof of Lemma* 5.3. Let us first prove two preliminary results stated in Lemmas 5.6 and 5.7.

In this section, ℓ^ν is related to ν through (2.13), $\mathbb{Q}^{\nu,n}$ is defined by (2.3) and $Q^{\nu,n}$ is the corresponding law of X^n. Let us derive an alternate form of the log-likelihood $\log\frac{dQ^{\nu,n}}{dP^n}$. For $\xi \in D_{\mathcal{P}}$, let $F_\xi(t,i) \stackrel{\Delta}{=} \sum_{j\le i}\xi_t(j)$, and let us define

$$\begin{aligned} I^{\nu,n}(\xi) \stackrel{\Delta}{=} {} & \sum_{i\ge 0} F_\xi(0,i)\log\ell^\nu_0(i) - \sum_{i\ge 0} F_\xi(T,i)\log\ell^\nu_T(i) \\ & + \sum_{k=1}^{\lfloor nT\rfloor -1}\frac{1}{n}\sum_{i\ge 0} F_\xi\Big(\frac{k}{n}, i\Big)\frac{\log\ell^\nu_{k/n}(i) - \log\ell^\nu_{(k-1)/n}(i)}{1/n} \\ & - \sum_{k=0}^{\lfloor nT\rfloor}\frac{1}{n}\log\langle \ell^\nu_{k/n}, \xi_{k/n}\rangle, \end{aligned} \tag{5.2}$$

$$\begin{aligned} I^{\nu}(\xi) \stackrel{\Delta}{=} {} & \sum_{i\ge 0} F_\xi(0,i)\log\ell^\nu_0(i) - \sum_{i\ge 0} F_\xi(T,i)\log\ell^\nu_T(i) \\ & + \int_{[0,T]}\Big[\sum_{i\ge 0} F_\xi(t,i)\partial_t\log\ell^\nu_t(i)\Big]\,dt - \int_{[0,T]}\log\langle \ell^\nu_t, \xi_t\rangle\,dt. \end{aligned} \tag{5.3}$$

LEMMA 5.6. *For any ν satisfying conditions* (i), (ii) *and* (iv) *in Definition* 5.2, $\frac{1}{n}\log\frac{dQ^{\nu,n}}{dP^n}(X^n) = I^{\nu,n}(X^n)$.

PROOF. The result follows from

$$\log\frac{dQ^{\nu,n}}{dP^n}(X^n) = \sum_{k=0}^{\lfloor nT\rfloor -1}\sum_{i=0}^{\infty}\mathbb{1}_{\{S^n_k(B^n_{k+1})=i\}}\log\bigg(\frac{\ell^\nu_{k/n}(i)}{\langle \ell^\nu_{k/n}, X^n_{k/n}\rangle}\bigg)$$

$$
\begin{aligned}
&\overset{(a)}{=} \sum_{k=0}^{\lfloor nT\rfloor-1} \sum_{i=0}^{\infty} \log \ell^{\nu}_{k/n}(i) \sum_{j\le i} n[X^n_{k/n}(j) - X^n_{(k+1)/n}(j)] \\
&\quad - \sum_{k=0}^{\lfloor nT\rfloor-1} \log\langle \ell^{\nu}_{k/n}, X^n_{k/n}\rangle \\
&\overset{(b)}{=} n \sum_{i\ge 0} \log \ell^{\nu}_0(i) \sum_{j\le i} X^n_0(j) - n \sum_{i\ge 0} \log \ell^{\nu}_{T-1/n}(i) \sum_{j\le i} X^n_T(j) \\
&\quad + n \sum_{k=1}^{\lfloor nT\rfloor-1} \frac{1}{n} \sum_{i\ge 0} \frac{\log \ell^{\nu}_{k/n}(i) - \log \ell^{\nu}_{(k-1)/n}(i)}{1/n} \sum_{j\le i} X^n_{k/n}(j) \\
&\quad - \sum_{k=0}^{\lfloor nT\rfloor-1} \log\langle \ell^{\nu}_{k/n}, X^n_{k/n}\rangle,
\end{aligned}
$$

where (a) comes from the identity $\mathbb{1}_{\{S^n_k(B^n_{k+1})=i\}} = n\sum_{j\le i}(X^n_{k/n}(j) - X^n_{(k+1)/n}(j))$, and (b) is Abel's transformation $\sum_{i=0}^n a_i(b_i - b_{i+1}) = \sum_{i=1}^n b_i(a_i - a_{i-1}) - b_{n+1}a_n$. $\square$

LEMMA 5.7. *For any ν satisfying conditions* (i), (ii) *and* (iv) *in Definition* 5.2, *we have $I^{\nu}(\nu) = I(\nu)$.*

PROOF. For any nice ν, we have

$$
\begin{aligned}
I^{\nu}(\nu) &\overset{(a)}{=} \sum_{i\ge 0}\Bigl[F_\nu(0,i)\log \ell^{\nu}_0(i) - F_\nu(T,i)\log \ell^{\nu}_T(i) + \int_{[0,T]} F_\nu(t,i)\partial_t \log \ell^{\nu}_t(i)\,dt\Bigr] \\
&\overset{(b)}{=} \sum_{i\ge 0} \int_{[0,T]} -\partial_t F_\nu(t,i) \log \ell^{\nu}_t(i)\,dt \\
&\overset{(c)}{=} \int_{[0,T]} \langle \ell^{\nu}_t \log \ell^{\nu}_t, \nu_t\rangle\,dt,
\end{aligned}
$$

where (a) follows from Fubini's theorem and $\log\langle \ell^{\nu}_t, \nu_t\rangle = 0$ for all t, by (ii) in Proposition 2.8, (b) follows from an integration by parts (ℓ^{ν} is t-differentiable), (c) is a consequence of the definition of ℓ^{ν} [see (2.13)] and an application of Fubini's theorem. The lemma then follows from Proposition 2.8. $\square$

PROOF OF LEMMA 5.3 (Nice ν's are really nice). Let us prove that for any $\varepsilon > 0$, there exists an open neighborhood V of ν such that

$$
\liminf_n \inf\{-I^{\nu,n}(\xi); \xi \in V\} \ge \inf\{-I^{\nu}(\xi); \xi \in V\} - \varepsilon. \tag{5.4}
$$

Recall the definition of $I^{\nu,n}$ (5.2). The third summand on the right-hand side may be decomposed as $A+B+C$ with

$$A = \sum_{k=1}^{\lfloor nT\rfloor -1} \frac{1}{n}\sum_i F_\xi\left(\frac{k}{n},i\right)\left[\frac{\log \ell^\nu_{k/n}(i) - \log \ell^\nu_{(k-1)/n}(i)}{1/n} - \partial_t \log \ell^\nu_{(k-1)/n}(i)\right],$$

$$B = \sum_{k=1}^{\lfloor nT\rfloor -1} \frac{1}{n}\sum_i \left[F_\xi\left(\frac{k}{n},i\right) - F_\nu\left(\frac{k}{n},i\right)\right]\partial_t \log \ell^\nu_{(k-1)/n}(i),$$

$$C = \sum_{k=1}^{\lfloor nT\rfloor -1} \frac{1}{n}\sum_i F_\nu\left(\frac{k}{n},i\right)\partial_t \log \ell^\nu_{(k-1)/n}(i).$$

By condition (iii) in Definition 5.2, index i in summation ranges between 0 and $M-1$. In this proof, K stands for a nonnegative constant that may vary from line to line.

Control of A. Thanks to (iv) in Definition 5.2,

$$|n(\log \ell^\nu_{k/n}(i) - \log \ell^\nu_{k-1/n}(i)) - \partial_t \log \ell^\nu_{k-1/n}(i)| \le K/n,$$

and as $|F_\xi| \le 1$, we get $|A| \le KM/n$.

Control of B. Thanks to (iv) in Definition 5.2, $\sup_{t,i} |\partial_t \log \ell^\nu_t(i)| \le K$, and it is possible to find an open neighborhood V of ν such that $\sup_{t,i} |F_\xi(t,i) - F_\nu(t,i)| \le \varepsilon$ for all ξ in V. Therefore, $|B| \le KM\varepsilon$.

Control of C. As a Riemann series [note that $\partial_t \log \ell^\nu_t$ is continuous thanks to (iv) in Definition 5.2], $\lim_n C = \int_0^T \sum_i F_\nu(t,i)\partial_t \log \ell^\nu_t(i)\, dt$.

In order to control the fourth summand of the right-hand side of $I^{\nu,n}(\xi)$, note that thanks to (iii) in Definition 5.2, for some $L < \infty$,

$$\left|-\frac{1}{n}\log\langle \ell^\nu_t, \xi_t\rangle + \frac{1}{n}\log\langle \ell^\nu_t, \nu_t\rangle\right| \le \frac{L}{n}\|\nu_t - \xi_t\|.$$

Therefore, for any $\varepsilon > 0$, there exists an open neighborhood V of ν such that

$$\sup_{\xi\in V}\left|\sum_{k=0}^{\lfloor nT\rfloor} \frac{1}{n}[\log\langle \ell^\nu_{k/n}, \xi_{k/n}\rangle - \log\langle \ell^\nu_{k/n}, \nu_{k/n}\rangle]\right| \le L\varepsilon. \tag{5.5}$$

Combining the above arguments we have proved (5.4).

Note that (5.5) with the continuity of $\xi \mapsto F_\xi$ implies that I^ν is continuous at ν. Therefore

$$\sup_V \inf\{-I^\nu(\xi); \xi \in V\} = -I^\nu(\nu). \tag{5.6}$$

Observing by Lemma 5.6 that $Q^{\nu,n}$ and P^n are mutually absolutely continuous measures and that

$$Q^{\nu,n}\text{-}\operatorname*{ess\,inf}_{\xi\in V}\left(\frac{1}{n}\log\frac{dP^n}{dQ^{\nu,n}}(\xi)\right) \ge \inf\{-I^{\nu,n}(\xi); \xi \in V\},$$

the lemma follows by combining this inequality, (5.4), (5.6) and Lemma 5.7. $\square$

5.5. *Proof of Lemma* 5.4 (Nice ν's are dense). The proof of Lemma 5.4 is postponed to after the proofs of two technical lemmas (5.8 and 5.9) concerning two parametrized regularization procedures; see (5.7) and (5.13).

The first regularization proceeds by time extension, mixing, and convolution by the following kernel:

$$\zeta^\varepsilon(s) \triangleq \begin{cases} \dfrac{2}{\varepsilon}\left(1-\dfrac{s}{\varepsilon}\right), & \text{for } 0 \le s \le \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

Remember that $\mathbf{p}_t$ denotes the Poisson distribution with parameter t.

As the convolution by the regularizing kernel ζ^ε depends on sample paths up to time $T+\varepsilon$, we first introduce a time-extension $\tilde{\nu}$ of ν. For any ν in $\mathcal{D}_I$, let $\tilde{\nu}$ be defined by:

$$\tilde{\nu} \triangleq \begin{cases} \tilde{\nu}_t = \nu_t, & \text{for } t \in [0,T], \\ \dfrac{d\tilde{\nu}_t}{dt}(i) = \tilde{\nu}_t(i-1) - \tilde{\nu}_t(i), & \text{for } t > T \text{ and } i \ge 0, \end{cases}$$

and for $\alpha \in [0,1]$, let $\nu_t^{\alpha,\varepsilon}$ be defined for all $t \ge 0$ by

$$\nu^{\alpha,\varepsilon} \triangleq \zeta^\varepsilon * \nu^\alpha \quad \text{where } \nu_t^\alpha \triangleq (1-\alpha)\tilde{\nu}_t + \alpha \mathbf{p}_t; \tag{5.7}$$

that is, $\nu_t^{\alpha,\varepsilon} = \int_0^\infty \zeta^\varepsilon(s)\nu_{t+s}^\alpha\, ds$.

The paths $\tilde{\nu}$, ν^α and $\nu^{\alpha,\varepsilon}$ belong to $D([0,\infty),\mathcal{P}(\mathbb{N}))$. Notice that the restriction of $\nu^{\alpha,\varepsilon}$ to $[0,T]$ satisfies condition (ii) in the Definition 5.2 of nice ν's and that $\nu_\cdot^{\alpha,\varepsilon}(i)$ and $\ell^{\nu_\cdot^{\alpha,\varepsilon}}(i)$ are infinitely differentiable with respect to t for all i [see (2.13)]. Let $\tilde{I}$ be defined by

$$\tilde{I}(\tilde{\nu}) \triangleq \int_{[0,\infty)} H(\upsilon_t^{\tilde{\nu}} \mid \tilde{\nu}_t)\, dt,$$

where $\upsilon_t^{\tilde{\nu}}$ is defined by (2.7). The following lemma summarizes the main properties of the regularized sample path $\nu^{\alpha,\varepsilon}$.

LEMMA 5.8. *If $I(\nu) < \infty$, the following statements hold*:

$$\tilde{I}(\nu^\alpha) \le I(\nu), \tag{5.8}$$

$$\tilde{I}\big(\tilde{\nu}(\cdot + s)\big) \le I(\nu), \qquad \forall\, s \ge 0, \tag{5.9}$$

$$\tilde{I}(\nu^{\alpha,\varepsilon}) \le \tilde{I}(\nu^\alpha), \tag{5.10}$$

$$\lim_{\alpha\downarrow 0,\varepsilon\downarrow 0} \sup_{t\le T} \|\nu_t^{\alpha,\varepsilon} - \nu_t\| = 0, \tag{5.11}$$

$$\lim_{\alpha\downarrow 0,\varepsilon\downarrow 0} \tilde{I}(\nu^{\alpha,\varepsilon}) = I(\nu). \tag{5.12}$$

PROOF. Note that as $\mathbf{p}_t(i) > 0$ for all $t > 0$, $i \in \mathbb{N}$, for all $\alpha > 0$ we have $\nu_t^\alpha(i) > 0$, $\nu_t^{\alpha,\varepsilon}(i) > 0$, and thus $\ell_t^{\alpha,\varepsilon}(i)$ is uniquely defined by (2.13).

As $\tilde{\nu}_t = \nu_t$ and $\ell_t^{\tilde{\nu}} = \ell_t^\nu$ for $t \leq T$ and $\ell_t^{\tilde{\nu}} = 1$ for $t > T$, we have

$$\tilde{I}(\tilde{\nu}) = I(\nu) + \int_{[T,\infty)} \langle 1 \log 1, \tilde{\nu}_s \rangle \, ds = I(\nu).$$

The convexity of $\tilde{I}$ implies the inequalities (5.8) and (5.10).

The same remarks imply that time shifting may only decrease the rate function, that is, (5.9).

The convergence (5.11) follows from

$$\begin{aligned}
&\sup_{0 \leq t \leq T} \|\nu_t^{\alpha,\varepsilon} - \nu_t\| \\
&\qquad = \sup_{0 \leq t \leq T} \left\| \int_0^\infty \big((1-\alpha)\tilde{\nu}_{t+s} + \alpha \mathbf{p}_{t+s}\big)\zeta_s^\varepsilon \, ds - \nu_t \right\| \\
&\qquad \leq \sup_{0 \leq t \leq T} (1-\alpha) \left\| \int_0^\infty (\tilde{\nu}_{t+s} - \tilde{\nu}_t)\zeta_s^\varepsilon \, ds \right\| + \alpha \left\| \int_0^\infty (\mathbf{p}_{t+s} - \tilde{\nu}_t)\zeta_s^\varepsilon \, ds \right\| \\
&\qquad \leq \sup_{0 \leq t \leq T} (1-\alpha) \left\| \int_0^\infty (\tilde{\nu}_{t+s} - \tilde{\nu}_t)\zeta_s^\varepsilon \, ds \right\| + 2\alpha \quad (\text{as } \|\tilde{\nu}_{t+s} - \tilde{\nu}_t\| \leq 2s) \\
&\qquad \leq \int_0^\varepsilon 2s\zeta_s^\varepsilon \, ds + 2\alpha \quad \text{by Corollary 4.3} \\
&\qquad \leq \varepsilon + 2\alpha.
\end{aligned}$$

Proof of (5.12). Now we identify $\nu^{\alpha,\varepsilon}$ with its restriction to $[0, T]$. Inequalities (5.8) and (5.10) imply that $I(\nu^{\alpha,\varepsilon}) \leq I(\nu)$ for all $\alpha, \varepsilon > 0$. On the other hand, as I is lower semicontinuous, with (5.11), we obtain $\liminf_{\alpha,\varepsilon} I(\nu^{\alpha,\varepsilon}) \geq I(\nu)$. Hence, $\lim_{\alpha,\varepsilon} I(\nu^{\alpha,\varepsilon}) = I(\nu)$. $\square$

The second regularization procedure operates directly on ℓ^ν. For any integer M, let us define Φ_M by $\Phi_M \ell_t^\nu(i) = \ell_t^\nu(i)$ for $i \leq M$, $t \geq 0$, and $\Phi_M \ell_t^\nu(i) = 1$ for all $i > M$ and $t \geq 0$. The associated sample path ν^M is defined by

$$\dot{\nu}_t^M(i) = [\Phi_M \ell_t^\nu(i-1)]\nu_t^M(i-1) - [\Phi_M \ell_t^\nu(i)]\nu_t^M(i). \tag{5.13}$$

LEMMA 5.9. *If $I(\nu) < \infty$ and $\ell_\bullet^\nu(i)$ is t-continuous for all $i \geq 0$, the following statements hold*:

$$\lim_{M \to \infty} \sup_{t \leq T} \|\nu_t - \nu_t^M\| = 0, \tag{5.14}$$

$$\lim_{M \to \infty} I(\nu^M) = I(\nu). \tag{5.15}$$

PROOF. Thanks to the t-continuity of $\ell_{\cdot}^{\nu}(i)$ for all i, $\sup_{t,i} \Phi_M \ell_t^{\nu}(i) < \infty$. By construction, for any $i \le M$, $\nu_t^M(i) = \nu_t(i)$ for all t. It follows that

$$I(\nu) - I(\nu^M) = \int_{[0,T]} \left[\sum_{i>M} \nu_t(i) \ell_t^{\nu}(i) \log \ell_t^{\nu}(i) \right] dt.$$

Letting M tend to infinity, by dominated convergence, the right-hand side vanishes and (5.15) is established. As

$$\sum_{i \ge 0} |\nu_t^M(i) - \nu_t(i)| = \sum_i \left| \int_{[0,t]} \big(\dot{\nu}_s^M(i) - \dot{\nu}(i)\big)\, ds \right|$$

by Proposition 2.8,

$$\begin{aligned}
&\sum_{i \ge 0} |\nu_t^M(i) - \nu_t(i)| \\
&\quad = \sum_{i>M} \left| \int_{[0,t]} \big(\nu_s^M(i-1) - \nu_s^M(i) - \ell_s^{\nu}(i-1)\nu_s(i-1) + \ell_s^{\nu}(i)\nu_s(i)\big)\, ds \right| \\
&\quad \le 2 \int_{[0,t]} \sum_{i>M} |\nu_s^M(i) - \nu_s(i)|\, ds + 2 \sum_{i>M} \left| \int_{[0,t]} [\ell_s^{\nu}(i) - 1]\nu_s(i) \right| ds \\
&\quad \le 2 \int_{[0,t]} \sum_{i \ge 0} |\nu_s^M(i) - \nu_s(i)|\, ds + h^M(t),
\end{aligned}$$

where $h^M(t) \triangleq 2 \sum_{i>M} \int_{[0,t]} (\ell_s^{\nu}(i) + 1)\nu_s(i)\, ds$. Applying Gronwall's lemma,

$$\sum_{i \ge 0} |\nu_t^M(i) - \nu_t(i)| \le h^M(t) + 2 \int_0^t h^M(s) e^{2(t-s)}\, ds.$$

However, according to Dini's lemma, a sequence of continuous functions decreasing pointwise toward 0 on the compact interval $[0, T]$ is also uniformly convergent. Therefore (5.14) follows from the fact that h^M decreases pointwise to 0 as M tends to ∞. □

PROOF OF LEMMA 5.4. Let ν be in $\mathcal{D}_I$. First apply the regularization (5.7). Then apply the second regularization (5.13) to $\nu^{\alpha,\varepsilon}$ for α, ε small enough. The resulting path is nice and the desired result follows from Lemmas 5.8 and 5.9. □

6. Stronger topologies. Should stronger topologies be considered? Recall that X_t^n is similar to the empirical measure of a Poisson random variable with parameter t. The latter satisfies the LDP with respect to the total variation distance with the rate function $H(\cdot \mid \mathbf{p}_t)$. It is not reasonable to think of test functions that could be larger than $i \log i$: the distribution $\nu(i) \propto 1/(i^2 \log^{2+\delta}(i))$ with

$\delta > 0$ has finite relative entropy with respect to any Poisson distribution although $\langle i \log^{1+\varepsilon}(i), \nu\rangle = \infty$ as soon as $\varepsilon \geq \delta$ (see [16] for approaches to extension of Sanov's theorem).

Let p be strictly larger than 1 and let $\mathcal{H}$ be the class of sequences defined by

$$\mathcal{H} = \{G = (G(i))_{i\in\mathbb{N}} : |G(0)| = 1 \text{ and for } i \geq 1, |DG(i)| \leq \log^{1/p}(i)\}.$$

If $G \in \mathcal{H}$ then $|G(i)| \leq i \log^{1/p}(i)$. In this section we consider the following metric on $D_{\mathcal{P}}$:

$$d_{\mathcal{H}}(\nu, \nu') \overset{\Delta}{=} \sup_{s\in[0,T]} \sup_{G\in\mathcal{H}} (\langle G, \nu_s\rangle - \langle G, \nu'_s\rangle). \tag{6.1}$$

THEOREM 6.1. *The sequence (X^n) satisfies the LDP with rate function I on $D_{\mathcal{P}}$ equipped with the topology defined by the metric $d_{\mathcal{H}}$.*

The proof of Theorem 6.1 proceeds according to the following steps. The compactness of the level sets of I under metric $d_{\mathcal{H}}$ and the exponential equivalence between (X^n) and the linearly interpolated process $(\widehat{X}^n)$ are established. Finally the exponential tightness of $(\widehat{X}^n)$ is established using an exponential martingale argument. The theorem follows from the inverse contraction principle [8], Theorem 4.2.4.

LEMMA 6.2. *I is a rate function under the topology induced by $d_{\mathcal{H}}$.*

PROOF. The convexity and the lower-semicontinuity of I still hold. It is enough to prove that the finiteness of $I(\nu)$ implies both an upper bound on the modulus of continuity under metric $d_{\mathcal{H}}$ and that there exists a compact set K_α of $\mathcal{P}(\mathbb{N})$ such that $\nu_t \in K_\alpha$.

Let $\nu \in D_{\mathcal{P}}$ be such that $I(\nu) < \infty$ and $q = p/(p-1)$. For any $G \in \mathcal{H}$,

$$\begin{aligned}
\langle G, \nu_t - \nu_s\rangle &\overset{(a)}{=} \int_{[s,t]} \langle \ell_u^\nu DG, \nu_u\rangle\, du = \int_{[0,T]} \langle \mathbb{1}_{[s,t]} DG \ell_u^\nu, \nu_u\rangle\, du \\
&\overset{(b)}{\leq} \left[\int_{[0,T]} \langle \mathbb{1}_{[s,t]}, \ell_u^\nu \nu_u\rangle\, du\right]^{1/q} \times \left[\int_{[0,T]} \langle |DG|^p, \ell_u^\nu \nu_u\rangle\, du\right]^{1/p} \\
&\overset{(d)}{\leq} |t-s|^{1/q} \times \int_{[0,T]} \langle v, \ell_u^\nu \nu_u\rangle\, du,
\end{aligned} \tag{6.2}$$

where we set $v(0) = 1$ and for $i \geq 1$, $v(i) = \log(i)$. Indeed, (a) follows from Proposition 2.8, (b) follows from Hölder's inequality and (c) from the definition of $\mathcal{H}$. Now

$$\begin{aligned}
\int_{[0,T]} \langle v, \ell_u^\nu \nu_u\rangle\, du &\overset{(a)}{\leq} \int_{[0,T]} \langle (i-1)_{i\in\mathbb{N}}, \nu_u\rangle\, du + \int_{[0,T]} \langle \ell_u^\nu \log \ell_u^\nu, \nu_u\rangle\, du \\
&\overset{(b)}{\leq} \langle (i)_{i\in\mathbb{N}}, \nu_0\rangle T + \frac{T^2}{2} + I(\nu),
\end{aligned} \tag{6.3}$$

where (a) follows from the application of Young's inequality in the duality between τ and τ^*, that is, $xy \leq \tau(x) + \tau^*(y)$, and (b) from Proposition 2.8 again. Combining inequalities (6.2) and (6.3), we get

$$\sup_{G \in \mathcal{H}} \langle G, \nu_t - \nu_s \rangle \leq \left(\frac{T^2}{2} + \langle i, \nu_0 \rangle T + I(\nu) \right)^{1/p} (t-s)^{1/q}. \tag{6.4}$$

To check the compact containment property under metric $d_{\mathcal{H}}$, it is enough to check that if $I(\nu) < \infty$ and $\langle \phi, \nu_0 \rangle < \infty$ where $\phi(i) = (i \log i)$, then

$$\langle (i \log i)_{i \in \mathbb{N}}, \nu_t \rangle \leq \langle i \log i, \nu_0 \rangle + I(\nu) + e\left(\frac{t^2}{2} + (1 + \langle i+1, \nu_0 \rangle) t \right).$$

As $I(\nu) < \infty$, by Proposition 2.8, $d\langle \phi, \nu_t \rangle / dt = \langle D\phi\, \ell_t^\nu, \nu_t \rangle$,

$$\begin{aligned} \frac{d\langle \phi, \nu_t \rangle}{dt} &\overset{(a)}{\leq} \sum \nu_t(i) \big(\ell_t^\nu(i) \log \ell_t^\nu(i) - \big(\ell_t^\nu(i) - 1 \big) \big) \\ &\qquad + \sum \nu_t(i) \big(e^{D\phi(i)} - D\phi(i) - 1 \big) + \sum \nu_t(i) D\phi(i) \\ &\overset{(b)}{\leq} \langle \phi(\ell_t^\nu), \nu_t \rangle + e\,(t + 1 + \langle i+1, \nu_0 \rangle), \end{aligned}$$

where (a) follows from Young's inequality, and (b) follows from $D\phi(i) \leq 1 + \log(i+1)$, $\langle \ell^\nu, \nu \rangle = 1$ and $\langle i+1, \nu_t \rangle = \langle i+1, \nu_0 \rangle$. Integration with respect to t finishes the proof. $\square$

In the sequel, $\widehat{X}^n$ denotes the linearly interpolated version of X^n,

$$\widehat{X}_t^n \triangleq X^n_{\lfloor nt \rfloor / n} + \left(t - \frac{\lfloor nt \rfloor}{n} \right) (X^n_{\lceil nt \rceil / n} - X^n_{\lfloor nt \rfloor / n}).$$

By Theorem 4.2.13 in [8], the exponential equivalence between X^n and $\widehat{X}^n$ warrants that $\widehat{X}^n$ satisfies the LDP with rate function I under the topology induced by the total variation distance.

LEMMA 6.3. *X^n and $\widehat{X}^n$ are exponentially equivalent under the topology induced by $d_{\mathcal{H}}$.*

Let us denote by $L^n(k) \triangleq S^n_{k-1}(B^n_k)$ the occupancy score at time $(k-1)/n$ of the bins where the kth allocation takes place at time k/n.

PROOF. As for each $G \in \mathcal{H}$,

$$\sum_{i \geq 0} G(i) \big[X^n_{k/n}(i) - X^n_{(k+1)/n}(i) \big] = \frac{1}{n} DG\big(L^n(k)\big) \leq \frac{1}{n} \log^{1/p}\big(L^n(k)\big),$$

we have

$$\mathbb{P}\Big[\sup_{t\le T}\sup_{G\in\mathcal{H}}\langle G, X_t^n - \widehat{X}_t^n\rangle \ge \eta\Big]$$
$$\le \mathbb{P}\Big[\sup_{k\le nT}\frac{1}{n}\log L^n(k)\ge\eta\Big]\le nT\,\mathbb{P}[L^n(\lfloor nT\rfloor)\ge e^{n\eta}].$$

However, at time $T = k/n$, ess-sup $L^n(k)$ is smaller than nT. Hence

$$\forall\,\eta, \forall\, n, \qquad e^{n\eta} > nT \quad\Rightarrow\quad \sup_{t\le T}\sup_{G\in\mathcal{H}}\langle G, X_t^n - \widehat{X}_t^n\rangle < \eta.$$

So for all $\eta > 0$,

$$\lim\frac{1}{n}\log\mathbb{P}\Big[\sup_{t\le T}\sup_{G\in\mathcal{H}}\langle G, X_t^n - \widehat{X}_t^n\rangle \ge \eta\Big] = -\infty. \qquad\square$$

It remains to show that $(\widehat{X}^n)$ also satisfies the LDP with the rate function I with respect to the topology defined by (6.1). By the inverse contraction principle [8], it is enough to check the exponential tightness.

LEMMA 6.4. *The sequence $(\widehat{X}^n)$ is exponentially tight under the topology induced by metric $d_{\mathcal{H}}$.*

We will use the following corollary of the Cauchy–Schwarz inequality. Let $q(i)$ be a probability on $\mathbb{N}$ having finite expectation. Then for some universal constant C,

$$\sum_{i\ge0} q(i)^{3/4} \le C\Big(\sum_{i\ge0} i\,q(i)\Big)^{1/2} < \infty. \tag{6.5}$$

PROOF. In view of Lemma 6.2, it is enough to check that for some $\alpha > 0$,

$$\limsup_n \frac{1}{n}\log\mathbb{P}\{I(\widehat{X}^n) > \alpha\} < 0.$$

Note first that $I(\widehat{X}^n) = \sum_{k=0}^{\lfloor nT\rfloor} -1/n\log X^n_{(k-1)/n}(L^n_k)$. Denote by Z_m the following quantity:

$$Z_m \triangleq \prod_{k=0}^{m}\frac{[\widehat{X}^n_{(k-1)/n}(L^n_k)]^{-1/4}}{\sum_{i\ge0}[\widehat{X}^n_{(k-1)/n}(i)]^{3/4}}.$$

One may check that Z_m is an $\mathcal{A}_m$-martingale. This entails

$$\mathbb{E}_{\mathbb{P}^n}\Bigg[\exp\Bigg(\frac{n}{4}I(\widehat{X}^n) - \sum_{k=1}^{\lfloor nt\rfloor}\log\Bigg[\sum_{i\ge0}\big(X^n_{(k-1)/n}(i)\big)^{3/4}\Bigg]\Bigg)\Bigg] = \mathbb{E}_{\mathbb{P}^n}[Z_1]. \tag{6.6}$$

Now

$$Z_1 = \frac{[X_0^n(L_1^n)]^{-1/4}}{\sum_i [X_0^n(i)]^{3/4}} \le [X_0^n(L_1^n)]^{-1/4} = [\nu_0(L_1^n)]^{-1/4},$$

thus using the initial remark,

$$\mathbb{E}_{\mathbb{P}^n}[Z_1] \le \sum_{i\ge 0} [\nu_0(i)]^{3/4} < \infty.$$

On the other hand, by (6.5) and recalling $\langle (i)_{i\in\mathbb{N}}, \widehat{X}_T^n\rangle \le \langle (i)_{i\in\mathbb{N}}, \widehat{X}_0^n\rangle + T$,

$$\sum_{i\ge 0} [\widehat{X}_T^n(i)]^{3/4} \le C(\langle i, \widehat{X}_0^n\rangle + T)^{1/2} \triangleq K,$$

thus,

$$Z_{\lfloor nT\rfloor} \ge K^{-\lfloor nT\rfloor} \prod_{k=1}^{\lfloor nT\rfloor} [\widehat{X}^n_{(k-1)/n}(L_k^n)]^{-1/4}. \tag{6.7}$$

Finally,

$$\begin{aligned}
\mathbb{P}\{I(\widehat{X}^n) \ge \alpha\} &= \mathbb{P}\left\{K^{-\lfloor nT\rfloor} \exp\left(\frac{n}{4} I(\widehat{X}^n)\right) \ge K^{-\lfloor nT\rfloor} \exp\left(\frac{n\alpha}{4}\right)\right\}\\
&\le \mathbb{P}\left\{Z_{\lfloor nT\rfloor} \ge K^{-\lfloor nT\rfloor} \exp\left(\frac{n\alpha}{4}\right)\right\}\\
&\le \mathbb{E}_{\mathbb{P}^n}[Z_1] K^{\lfloor nT\rfloor} \exp\left(-\frac{n\alpha}{4}\right).
\end{aligned}$$

As $\mathbb{E}_{\mathbb{P}^n} Z_1 < \infty$,

$$\limsup_n \frac{1}{n} \log \mathbb{P}\{I(\widehat{X}^n) > \alpha\} \le -\left[-T\log K + \frac{\alpha}{4}\right],$$

which is negative for sufficiently large α. $\square$

7. An application to random graphs. Theorem 2.7 is used to characterize the large deviations of the degree sequence of sparse random graphs.

In the Erdös–Rényi $\mathcal{G}(n, \lfloor tn\rfloor)$ model for random graphs, $\lfloor tn\rfloor$ edges are inserted at random among n vertices. When t remains fixed while n tends to infinity, the model deals with sparse random graphs (with average degree $2t$). The degree of vertex i after $k = \lfloor nt\rfloor$ edge insertions is denoted $U_i^n(k)$. Any (random) graph defines an empirical probability measure V_t^n on $\mathbb{N}$:

$$V_t^n \triangleq \frac{1}{n}\sum_{i=1}^n \delta_{U_i^n(\lfloor tn\rfloor)},$$

which is called the *degree distribution of the graph* [3]. If vertices are identified with bins and edge extremities with balls, the degree distribution may be viewed as a conditioned empirical occupancy measure [20, 4, 6]. The conditioning approach is fruitful when establishing upper bounds, but it runs into difficulties when trying to prove LDP. Hence, we will depend on coupling and exponential approximation arguments to derive an LDP for the degree distribution of sparse random graphs.

THEOREM 7.1. *In the $\mathcal{G}(n, \lfloor tn \rfloor)$ random graph model, the empirical degree distribution satisfies a LDP with the rate function*

$$I'(\mu) \stackrel{\Delta}{=} \inf\{I(\nu) : \nu \in D_{\mathcal{P}}([0, 2t], \mathcal{P}(\mathbb{N})), \nu_{2t} = \mu\}.$$

The theorem follows from the coupling lemma below and Theorem 4.2.13 in [8].

LEMMA 7.2. *There exists a sequence of probability spaces over which one may define a random variable $(Y_t^n)_{t \le T}$ distributed like (X_t^n) and another random variable $(W_t^n)_{t \le T}$ distributed like $(V_t^n)_{t \le T}$ and such that for any $\varepsilon > 0$,*

$$\lim_n \frac{1}{n} \log \mathbb{P}\{\sup_t \|Y_{2t}^n - W_t^n\| > \varepsilon\} = -\infty.$$

PROOF. The coupling space is defined as follows. After step k, $2k$ balls have been inserted into the n bins and k edges have been inserted among the n vertices. At step $k+1$, a couple of indices (i, i') is picked uniformly at random, a ball is inserted into bin i and another ball is inserted into bin i' (both bins may be identical). If $i \neq i'$ and if the edge $\{i, i'\}$ had not been inserted previously then the edge $\{i, i'\}$ is inserted; otherwise a new couple of indices is picked at random until the couple defines a new edge, then this edge is inserted into the random graph under construction.

Notice that the probability that the pair of bins that receive the two balls at step k differs from the pair of vertices adjacent to the kth edge is equal to $\frac{1}{n} + (1 - \frac{1}{n})\frac{2k}{n(n-1)} \le \frac{1}{n}(1 + 2T)$. Let Δ^n denote the total number of steps with index less than nT at which the insertion in the random allocation process and the insertion in the graph construction process differ. It is worth noting that

$$\sup_{t \le T} \|Y_{2t}^n - W_t^n\| \le \frac{8\Delta^n}{n}. \tag{7.1}$$

As S^n is a sum of independent Bernoulli random variables with success probability $(\frac{1}{n} + (1 - \frac{1}{n})\frac{2k}{n(n-1)})_{k \le nT}$, $\mathbb{E}\Delta^n \le T(1+T)$ and $\mathrm{Var}(\Delta^n) \le T(1+2T)$. Now applying Bernstein's inequality, we get

$$\mathbb{P}\{\Delta^n \ge T(1+T) + s\} \le \exp\left[-\frac{s^2}{2(T(1+2T) + s/3)}\right]. \tag{7.2}$$

The lemma follows by combining inequalities (7.2) and (7.1). $\square$

For sparse random graphs, Theorem 7.1 complements the results reported in [17] where events with polynomially small probability are characterized.

Acknowledgment. The authors thank a careful and patient referee for providing many suggestions.

REFERENCES

[1] BARBOUR, A., HOLST, L. and JANSON, S. (1992). *Poisson Approximation.* Clarendon Press, London.

[2] BARTLETT, M. S. (1938). The characteristic function of a conditional statistic. *J. London Math. Soc.* **13** 62–67.

[3] BOLLOBÀS, B. (1985). *Random Graphs.* Academic Press, New York.

[4] BOLLOBÀS, B. and FRIEZE, A. M. (1985). On matching and Hamiltonian cycles in random graphs. In *Random Graphs 83. Annals of Discrete Mathematics* **28** 23–46.

[5] BOUCHERON, S. and GARDY, D. (1997). An urn model from learning theory. *Random Structures Algorithms* **10** 43–67.

[6] CHVATAL, V. (1991). Almost all graphs with 1.44 edges are 3 colourable. *Random Structures Algorithms* **2** 11–28.

[7] DEMBO, A., VERSHIK, A. and ZEITOUNI, O. (2000). Large deviations for integer partitions. *Markov Process. Related Fields* **6** 147–179.

[8] DEMBO, A. and ZEITOUNI, O. (1999). *Large Deviation Techniques and Applications.* Springer, New York.

[9] ETHIER, S. and KURTZ, T. G. (1986). *Markov Processes, Characterization and Convergence.* Wiley, New York.

[10] FELLER, W. (1968). *Probability Theory.* Wiley, New York.

[11] GILES, J. R. (1992). Convex analysis with application in the differentiation of convex functions. In *Research Notes in Math.* **58**. Pitman, London.

[12] JOHNSON, N. L. and KOTZ, S. (1977). *Urn Models and Their Application. An Approach to Modern Discrete Probability Theory.* Wiley, New York.

[13] KOLCIN, V. F., SEVAST'JANOV, B. A. and CISTJAKOV, V. P. (1978). *Random Allocations.* Wiley, New York.

[14] LÉONARD, C. (1995). Large deviations for particle systems associated with spatially homogeneous Boltzmann type equations. *Probab. Theory Related Fields* **101** 1–44.

[15] LÉONARD, C. (2000). Minimizers of energy functionals under not very integrable constraints. Preprint.

[16] LÉONARD, C. and NAJIM, J. (2000). An extension of Sanov's theorem. Application to the Gibbs conditioning principle. Preprint.

[17] MCKAY, B. D. and WORMALD, N. C. (1998). The degree sequence of a random graph. *Random Structures Algorithms* **11** 97–117.

[18] MORRIS, C. (1975). Central Limit Theorems for multinomial sums. *Ann. Statist.* **3** 165–168.

[19] PALKA, Z. (1984). On the number of vertices of given degree in a random graph. *J. Graph Theory* **8** 267–270.

[20] PITTEL, B., SPENCER, J. and WORMALD, N. (1996). Sudden emergence of a giant k-core in a random graph. *J. Combin. Theory Ser. B* **67** 111–151.

[21] QUINE, M. P. (1979). A functional Central Limit Theorem for generalized occupancy numbers. *Stochastic Process. Appl.* **9** 109–115.

[22] QUINE, M. P. and ROBINSON, J. (1982). A Berry–Essen bound for an occupancy problem. *Ann. Probab.* **10** 663–671.

[23] QUINE, M. P. and ROBINSON, J. (1984). Normal approximations to sums of scores based on occupancy numbers. *Ann. Probab.* **12** 794–804.

[24] RAO, R. and REN, R. (1991). *Theory of Orlicz Spaces*. Marcel Dekker, New York.

[25] ROCKAFELLAR, R. T. (1968). Integrals which are convex functionals. *Pacific J. Math.* **24** 525–538.

[26] SION, M. (1958). On general Minimax theorems. *Pacific J. Math.* **8** 171–176.

[27] VARADHAN, S. R. S. (1984). *Large Deviations and Applications*. SIAM, Philadelphia.

[28] VITTER, J. S. and FLAJOLET, P. (1990). Average-case analysis of algorithms and data structures. In *Handbook of Theoretical Computer Science* **A** (J. Van Leeuwen, ed.) 431–524. Elsevier, Amsterdam.

LRI UMR CNRS 8623
UNIVERSITÉ PARIS-SUD
91405 ORSAY
FRANCE
E-MAIL: bouchero@lri.fr

LSP UMR C5583
118 ROUTE DE NARBONNE
UNIVERSITÉ PAUL SABATIER
31062 TOULOUSE
FRANCE
E-MAIL: gamboa@cict.fr

MODAL-X
UNIVERSITÉ PARIS 10, BÂT. G
200 AVENUE DE LA RÉPUBLIQUE
92001 NANTERRE
FRANCE
AND ECOLE POLYTECHNIQUE, CMAP
91128 PALAISEAU
FRANCE
E-MAIL: christian.leonard@u-paris10.fr