Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Stochastic Processes and their Applications 121 (2011) 2272–2302
stochastic processes and their applications
www.elsevier.com/locate/spa

Large deviations for the local fluctuations of random walks

Julien Barral[a,*], Patrick Loiseau[b,1]

[a] *LAGA (UMR 7539), Département de Mathématiques, Institut Galilée, Université Paris 13, 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse, France*
[b] *INRIA Paris-Rocquencourt, Domaine de Voluceau, BP 105, 78153 Le Chesnay Cedex, France*

Received 10 February 2011; received in revised form 8 June 2011; accepted 16 June 2011
Available online 2 July 2011

Abstract

We establish large deviation properties valid for almost every sample path of a class of stationary mixing processes $(X_1, \dots, X_n, \dots)$. These properties are inherited from those of $S_n = \sum_{i=1}^n X_i$ and describe how the local fluctuations of almost every realization of S_n deviate from the almost sure behavior. These results apply to the fluctuations of Brownian motion, Birkhoff averages on hyperbolic dynamics, as well as branching random walks. Also, they lead to new insights into the "randomness" of the digits of expansions in integer bases of Pi. We formulate a new conjecture, supported by numerical experiments, implying the normality of Pi.

Keywords: Large deviations; Random walks; Mixing processes; Hyperbolic dynamics; Random coverings; Normal numbers

1. Introduction

Given a sequence of i.i.d. real valued random variables $(X_n)_{n\geq 1}$, large deviations theory provides a precise estimate of the probability that the random walk $S_n = \sum_{i=1}^n X_i$ deviates from its almost sure asymptotic behavior, as long as X possesses finite exponential moments on

* Corresponding author.
E-mail addresses: barral@math.univ-paris13.fr (J. Barral), ploiseau@soe.ucsc.edu (P. Loiseau).
[1] Current address: University of California, School of Engineering, SOE-3, 1156 High Street, Santa Cruz, CA 95064, USA.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2011.06.004

a nontrivial domain. In particular, if $\Lambda(\lambda) = \log \mathbb{E}(\exp(\lambda X_1))$ is finite over an interval $\mathcal{D}_\Lambda$ whose interior contains 0, then (see Cramer's theorem in [11] Ch 2.2 or Gärtner–Ellis' theorem at the end of this section)

$$\forall\, x \in \Lambda'(\mathring{\mathcal{D}}_\Lambda), \quad \lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log\mu_n([x-\epsilon, x+\epsilon]) = -\Lambda^*(x), \tag{1.1}$$

where μ_n is the distribution of S_n/n and $\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}} \lambda x - \Lambda(\lambda)$. Notice that the case $x = \Lambda'(0) = \mathbb{E}(X_1)$ corresponds to the almost sure asymptotic behavior of $S_n(\omega)$ given by the strong law of large numbers: $S_n(\omega)/n \to \mathbb{E}(X_1)$ as $n \to \infty$, and $\Lambda^*(\mathbb{E}(X_1)) = 0$.

In this paper, we show that this large deviation principle (LDP) is transferred to almost every path of the random walk, though the behavior of $S_N(\omega)/N$ is prescribed by the strong law of large numbers. To see this, we look at the deviations from this behavior over the blocks $(X_{(j-1)n+1}(\omega), \dots, X_{jn}(\omega))$ of length $n \ll N$ picked up in $(X_i(\omega))_{1\le i\le N}$. Specifically, we define

$$\Delta S_n(j,\omega) = S_{jn}(\omega) - S_{(j-1)n}(\omega)$$

and for $N = k(n)\cdot n$ with $k(n) \to \infty$ as $n \to \infty$, we seek a LDP providing the almost sure asymptotic behavior of $\#\{1 \le j \le k(n) : \Delta S_n(j,\omega) \in [n(x-\epsilon), n(x+\epsilon)]\}$ and its possible connection with (1.1). Such a LDP would describe the local fluctuations of S_n. We shall obtain the following result as a special case of a more general statement (Theorem 2.3). We consider the random sequence of Borel measures $(\mu_n^\omega)_{n\ge1}$ on $\mathbb{R}$ defined as

$$\mu_n^\omega(B) = \frac{\#\{1 \le j \le k(n) : \Delta S_n(j,\omega)/n \in B\}}{k(n)} \quad \text{(for every Borel set } B),$$

as well as their logarithmic generating functions

$$\Lambda_n^\omega(\lambda) = \frac{1}{n}\log\int_{\mathbb{R}}\exp(n\lambda x)\,\mathrm{d}\mu_n^\omega(x) = \frac{1}{n}\log\left(\frac{1}{k(n)}\sum_{j=1}^{k(n)}\exp(\lambda\Delta S_n(j,\omega))\right) \quad (\lambda\in\mathbb{R}).$$

Theorem 1.1. *Let $(k(n))_{n\ge1}$ be a sequence of positive integers. Let $\lambda_0 \in \mathring{\mathcal{D}}_\Lambda$ and denote $\Lambda'(\lambda_0)$ as x_0.*

(1) *If $\liminf_{n\to\infty}\frac{\log k(n)}{n} > \Lambda^*(x_0)$ then there exists a neighborhood U of λ_0 in $\mathring{\mathcal{D}}_\Lambda$ such that, with probability 1, for all $\lambda \in U$*

$$\lim_{n\to\infty}\Lambda_n^\omega(\lambda) = \Lambda(\lambda), \tag{1.2}$$

and

$$\lim_{\varepsilon\to0}\lim_{n\to\infty}\frac{1}{n}\log\mu_n^\omega([x_0-\epsilon, x_0+\epsilon]) = -\Lambda^*(x_0).$$

(2) *If $\limsup_{n\to\infty}\frac{\log k(n)}{n} < \Lambda^*(x_0)$ and ϵ is small enough, with probability 1, for n large enough the set $\{1 \le j \le k(n) : \Delta S_n(j,\omega) \in [n(x_0-\epsilon), n(x_0+\epsilon)]\}$ is empty.*

(3) *If $\lim_{n\to\infty}\frac{\log k(n)}{n} = \Lambda^*(x_0)$ then, with probability 1, for all $t \ge 1$ we have*

$$\lim_{n\to\infty}\Lambda_n^\omega(t\lambda_0) = \Lambda(\lambda_0) + (t-1)\lambda_0 x_0. \tag{1.3}$$

Remark 1.1. (1) The previous result will be extended to weakly dependent sequences such that $\Lambda(\lambda) = \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}(\exp(\lambda S_n))$ converges as n tend to ∞ for each λ in an open interval.

For such sequences, one also has a strong law of large numbers so that for each $n_0 \geq 1$, one has

$$\lim_{k\to\infty} \frac{1}{n_0} \log\left(\frac{1}{k}\sum_{j=1}^{k} \exp(\lambda \Delta S_{n_0}(j,\omega))\right) = \frac{1}{n_0}\log \mathbb{E}(\exp(\lambda S_{n_0})),$$

hence

$$\lim_{n\to\infty}\lim_{k\to\infty} \frac{1}{n} \log\left(\frac{1}{k}\sum_{j=1}^{k} \exp(\lambda \Delta S_{n}(j,\omega))\right) = \Lambda(\lambda). \tag{1.4}$$

In Theorem 2.5 we give, in terms of the growth of $\log(k)/n$, a fine measurement of how $\frac{1}{n}\log\left(\frac{1}{k}\sum_{j=1}^{k}\exp(\lambda\Delta S_n(j,\omega))\right)$ is close to $\Lambda(\lambda)$.

(2) Theorem 1.1 cannot be obtained as a consequence of (1.4).

Let us show how Theorem 1.1 applies to the description of the dyadic expansion of real numbers. For $t \in [0,1]$ and $i \geq 1$ denote by t_i the ith digit of the dyadic expansion of t (the dyadic points, which have two expansions, are of no influence in our study): $t = \sum_{i\geq 1} t_i 2^{-i}$. Let $\mathbb{P}_p$ stand for the Bernoulli product of parameter $p \in (0,1)$, so that the $X_i(t) = t_i$ are i.i.d. Bernoulli variables of parameter p under $\mathbb{P}_p$ ($\mathbb{P}_{1/2}$ is the Lebesgue measure). By the strong law of large numbers, for $\mathbb{P}_p$-almost every t, $\lim_{N\to\infty}\sum_{i=1}^{N} t_i/N = p$. Here, $\mathcal{D} = \mathbb{R}$, $\Lambda(\lambda) = \log(1-p+p\exp(\lambda))$, $\Lambda'(\mathbb{R}) = (0,1)$, and $\Lambda^*(x) = x\log(x/p) + (1-x)\log((1-x)/(1-p)) = H(\mathbb{P}_x|\mathbb{P}_p)$ for all $x \in (0,1)$. As a consequence of Theorem 1.1(1), if $\liminf_{n\to\infty} \frac{\log k(n)}{n} > -\min(\log(p), \log(1-p))$, for $\mathbb{P}_p$-almost every t, for all $x \in (0,1)$

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} \frac{1}{n}\log\left(\frac{1}{k(n)}\#\left\{1 \leq j \leq k(n) : \left(\sum_{(j-1)n < i \leq jn} t_i\right) \in [n(x-\epsilon), n(x+\epsilon)]\right\}\right)$$
$$= -\Lambda^*(x).$$

Once one has such a result, it is very tempting to investigate whether or not it highlights questions related to the distribution of digits for numbers suspected to be normal in a given integer basis $m \geq 2$, i.e. such that for every $n_0 \geq 1$, for every finite sequence $(\varepsilon_1, \ldots, \varepsilon_{n_0}) \in \{0, \ldots, m-1\}^{n_0}$, the frequency of the occurrence of $(\varepsilon_1, \ldots, \varepsilon_{n_0})$ in the m-adic expansion of $t = \sum_{i\geq 1} t_i m^{-i}$ is equal to m^{-n_0}, i.e.

$$\lim_{k\to\infty}\frac{1}{k}\#\{1 \leq i \leq k : (t_i, \ldots, t_{i+n_0-1}) = (\varepsilon_1 \cdots \varepsilon_{n_0})\} = m^{-n_0}. \tag{1.5}$$

Indeed, for such numbers like the fractional part of Pi, numerical experiments support the conjecture that (1.5) holds, showing that these numbers should share statistical properties with almost every realization of a sequence X of independent random variables uniformly distributed in $\{0, \ldots, m-1\}$, and in this sense are "random". The recent discovery of the so-called BBP algorithm [1], to compute the nth digit without computing the preceding digits, has opened new perspectives on this question [2]. Theorem 1.1 leads to strengthen the conjecture about the "randomness" of Pi: the sequence of digits of Pi in a given integer basis obeys the same large deviation properties as almost every realization of X (see Conjecture 4.1 for a precise statement). This conjecture, which implies the normality of Pi, is supported by numerical experiments presented in Section 4.

We will obtain extensions of Theorem 1.1 valid for a class of $\mathbb{R}^d$-valued stationary mixing processes. We will also obtain a general result concerning the transfer of LDPs valid for random walks taking values in a separable normed vector space to LDPs valid for the local fluctuations of almost every realization of such random walks. These results are stated in Section 2 and illustrated with several natural examples in Section 3, namely Brownian motion, Birkhoff sums on symbolic spaces and some of their geometric realizations, branching random walks on Galton–Watson trees, and Poissonian random walks on Poisson point processes. The proofs of the main results are given in Sections 5 and 6.

We end this section by recalling general facts about large deviations theory.

General facts about large deviations theory

Let $\mathcal{Y}$ be a topological space and $B_{\mathcal{Y}}$ stand for the completed Borel σ-field. Let $(\mu_n)_{n\geq 1}$ be a sequence of probability measures on $(\mathcal{Y}, B_{\mathcal{Y}})$. Let $I : \mathcal{Y} \to [0, \infty]$ be a lower semi-continuous function. The domain of I is defined as $\mathcal{D}_I = \{x : I(x) < \infty\}$.

One says (see [11] Ch. 1.2) that $(\mu_n)_{n\geq 1}$ satisfies in $\mathcal{Y}$ the LDP with rate function I if for all set $\Gamma \in B_{\mathcal{Y}}$.

$$-\inf_{x\in\mathring{\Gamma}} I(x) \leq \liminf_{n\to\infty} \frac{1}{n}\log\mu_n(\Gamma) \leq \limsup_{n\to\infty}\frac{1}{n}\log\mu_n(\Gamma) \leq -\inf_{x\in\overline{\Gamma}} I(x). \tag{1.6}$$

The function I is said to be a good rate function if, moreover, for any $\alpha \in \mathbb{R}_+$ the level set $\{x \in \mathcal{Y} : I(x) \leq \alpha\}$ is compact.

One says that $(\mu_n)_{n\geq 1}$ satisfies in $\mathcal{Y}$ the weak LDP with rate function I if the upper bound in (1.6) holds when Γ is a compact subset of $\mathcal{Y}$.

The sequence $(\mu_n)_{n\geq 1}$ is said to be exponentially tight if for every $\alpha < \infty$ there exists a compact set $K_\alpha \subset \mathcal{Y}$ such that $\limsup_{n\to\infty} \frac{1}{n}\log\mu_n(K_\alpha^c) \leq -\alpha$. In this case, if $(\mu_n)_{n\geq 1}$ satisfies the weak LDP with rate function I then it satisfies the LDP with good rate function I (see [11] Lemma 1.2.18).

Large deviation principles have been derived successfully for various stochastic processes (see, e.g. [13,40,42,11]) as well as for dynamical systems (see, e.g. [31,32,44]).

The Gärtner–Ellis theorem

It is sometimes possible to derive, or relate, such a principle with the logarithmic generating functions of the measures μ_n whenever $\mathcal{Y}$ is a topological vector space. In this paper, when we use such a connection, we take $\mathcal{Y} = \mathbb{R}^d$ ($d \geq 1$). Then, the main tool is the Gärtner–Ellis theorem whose statement requires the following assumptions and definitions (see [11] Ch. 2.3, and [11] Ch. 4.5.3 for a version in topological vector spaces). Let $\langle\cdot,\cdot\rangle$ stand for the canonical scalar product on $\mathbb{R}^d$. Let $(\mu_n)_{n\geq 1}$ be a sequence of probability measures on $(\mathbb{R}^d, B_{\mathbb{R}^d})$. For each $n \geq 1$ let

$$\Lambda_n(\lambda) = \frac{1}{n}\log\int_{\mathbb{R}^d}\exp(n\langle\lambda, x\rangle)\,\mathrm{d}\mu_n(x) \quad (\lambda \in \mathbb{R}^d).$$

Assume (**A**): For each $\lambda \in \mathbb{R}^d$, $\Lambda(\lambda) = \lim_{n\to\infty}\Lambda_n(\lambda)$ exists as an extended real number. Further, the origin belongs to the interior of $\mathcal{D}_\Lambda = \{\lambda \in \mathbb{R}^d : \Lambda(\lambda) < \infty\}$.

The Fenchel–Legendre transform of Λ is defined as

$$\Lambda^*(x) = \sup\{\langle\lambda, x\rangle - \Lambda(\lambda) : \lambda \in \mathbb{R}^d\} \quad (x \in \mathbb{R}^d),$$

and one sets $\mathcal{D}_{\Lambda^*} = \{x \in \mathbb{R}^d : \Lambda^*(x) < \infty\}$.

Definition 1.2. $y \in \mathbb{R}^d$ is an exposed point of Λ^* if for some $\lambda \in \mathbb{R}^d$ and all $x \neq y \in \mathbb{R}^d$, $\langle \lambda, y \rangle - \Lambda^*(y) > \langle \lambda, x \rangle - \Lambda^*(x)$. Such a λ is called an exposing hyperplane.

Definition 1.3. Let $L : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ and $\mathcal{D}_L = \{\lambda \in \mathbb{R}^d : L(\lambda) < \infty\}$. The function L is said essentially smooth if:

(a) $\mathring{\mathcal{D}}_L$ is non-empty.
(b) L is differentiable throughout $\mathring{\mathcal{D}}_L$.
(c) L is steep, namely, $\lim_{n\to\infty} |\nabla L(\lambda_n)| = \infty$ whenever $(\lambda_n)_{n\geq 1}$ is a sequence in $\mathring{\mathcal{D}}_L$ converging to a boundary point of $\mathring{\mathcal{D}}_L$.

Remark 1.2. Corollary 25.1.2 of [34] ensures that the exposed points of Λ^* are precisely those x of the form $\nabla\Lambda(\lambda)$ for some $\lambda \in \mathring{\mathcal{D}}_\Lambda$. Moreover, Theorem 25.5 of [34] ensures that Λ is differentiable almost everywhere in $\lambda \in \mathring{\mathcal{D}}_\Lambda$, and if it is differentiable everywhere in $\lambda \in \mathring{\mathcal{D}}_\Lambda$, then it is C^1.

Theorem 1.4 (*Gärtner–Ellis*)**.** *Under the above assumption* (**A**)*:*

(1) *For any closed set $F \subset \mathbb{R}^d$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq - \inf_{x\in F} \Lambda^*(x).$$

(2) *For any open set $G \subset \mathbb{R}^d$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(G) \geq - \inf_{x\in G\cap\mathcal{F}} \Lambda^*(x),$$

where $\mathcal{F}$ is the set of exposed points of Λ^* whose exposing hyperplane belongs to $\mathring{\mathcal{D}}_\Lambda$.

(3) *If Λ is essentially smooth and lower semi-continuous, then the LDP with good rate function Λ^* holds in $\mathbb{R}^d$ for $(\mu_n)_{n\geq 1}$.*

A local version of this result is the following. It can be deduced from the proof of Gärtner–Ellis' theorem. Throughout, $B(\lambda, r)$ stands for the closed ball of center λ and radius r.

Theorem 1.5. *Suppose that $\Lambda = \lim_{n\to\infty} \Lambda_n$ exists and is finite over an open set $\mathcal{D}$. At any point λ of $\mathcal{D}$ at which Λ is differentiable, one has*

$$\lim_{\epsilon\to 0} \lim_{n\to\infty} \frac{1}{n} \log \mu_n(B(\nabla\Lambda(\lambda), \epsilon)) = -\Lambda^*(\nabla\Lambda(\lambda)).$$

Remark 1.3. In Theorem 1.5 we do not require that $0 \in \mathcal{D}$. This is because one goes back to this assumption by a standard reduction, systematically used in the proof of the lower bound part of Gärtner–Ellis' theorem, as follows. Fix any $\lambda_0 \in \mathcal{D}$ and replace μ_n by the measure $\mathrm{d}\widetilde{\mu}_n(x) = \exp(n\langle \lambda_0, x\rangle - n\Lambda_n(\lambda_0))\mathrm{d}\mu_n(x)$. Then replace $\Lambda_n(\lambda)$ by $\widetilde{\Lambda}_n(\lambda) = \frac{1}{n} \log \int_{\mathbb{R}^d} \exp(n\langle \lambda, x\rangle)\, \mathrm{d}\widetilde{\mu}_n(x) = \Lambda_n(\lambda + \lambda_0) - \Lambda_n(\lambda_0)$.

2. Large deviation principle for local fluctuations of random walks

We need a few notation and definitions related to the notion of weak dependence.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Given two sub-σ-algebras $\mathcal{U}$ and $\mathcal{V}$ of $\mathcal{A}$, their α-mixing coefficient is defined as (see [33] for a detailed account):

$$\alpha(\mathcal{U}, \mathcal{V}) = \sup\{|\mathbb{P}(U)\mathbb{P}(V) - \mathbb{P}(U \cap V)| : U \in \mathcal{U}, V \in \mathcal{V}\}. \tag{2.1}$$

Let $(E, \mathcal{T})$ be a measurable space. We consider $X = (X_i)_{i\in\mathbb{N}_+}$, a stationary process defined on $(\Omega, \mathcal{A}, \mathbb{P})$ and taking values in E.

For each $n \geq 1$, we define

$$X^{(n)} = (X_i^{(n)})_{i\geq 1} = ((X_{n(i-1)+1}, \dots, X_{in}))_{i\geq 1}. \tag{2.2}$$

Let T denote the shift operation on $E^{\mathbb{N}_+}$:

$$T(x_1, x_2, \dots) = (x_2, \dots).$$

We will assume that X satisfies some mixing properties.

The mixing coefficients $(\alpha_{X,m})_{m\geq 0}$ of the sequence $(X_i)_{i\in\mathbb{N}_+}$ are defined as :

$$\alpha_{X,0} = 1/2 \text{ and } \alpha_{X,m} = \sup\{\alpha(\sigma(X_i), \sigma(X_{i+j})) : i \geq 1, j \geq m\} \quad \text{for } m \geq 1. \tag{2.3}$$

Then, for $u \in [0, 1]$, one defines

$$\alpha_X^{-1}(u) = \inf\{m : \alpha_{X,m} \leq u\}.$$

For each $n \geq 1$ we let $S_n\Phi(X)$ stand for a measurable function of $(X, TX, \dots, T^{n-1}X)$ taking values in a normed vector space $(\mathcal{Y}, \| \|)$ endowed with the completed Borel σ-field $B_{\mathcal{Y}}$. A typical example will be the Birkhoff sums $\sum_{j=0}^{n-1} \Phi(T^j X)$ associated with a measurable function $\Phi : E^{\mathbb{N}_+} \to \mathcal{Y}$.

For each $n \geq 1$, denote by μ_n the distribution of the random variable $S_n\Phi(X)/n$ (viewed under $\mathbb{P}$).

If $\mathcal{Y} = \mathbb{R}^d$, we define the sequence of logarithmic moment generating functions

$$\Lambda_n(\lambda) = \frac{1}{n} \log \mathbb{E} \exp(\langle\lambda, S_n\Phi(X)\rangle) = \frac{1}{n} \log \int_{\mathbb{R}^d} \exp(n\langle\lambda, x\rangle)\, \mathrm{d}\mu_n(x) \quad (\lambda \in \mathbb{R}^d). \tag{2.4}$$

Our results will use assumptions among the following. They are divided into three types.

(1) Large deviation properties.

(**A1**) The sequence $(\mu_n)_{n\geq 1}$ satisfies in $\mathcal{Y}$ the LDP with rate function denoted by I.

(**A1**′) $\mathcal{Y} = \mathbb{R}^d$, and there exists a convex open set $\mathcal{D}$ in $\mathbb{R}^d$ such that

$$\Lambda(\lambda) = \lim_{n\to\infty} \Lambda_n(\lambda)$$

exists and is finite for each $\lambda \in \mathcal{D}$.

(**A1**″) $\mathcal{Y} = \mathbb{R}^d$, and for each $\lambda \in \mathbb{R}^d$, $\Lambda(\lambda) = \lim_{n\to\infty} \Lambda_n(\lambda)$ exists as an extended real number, and the origin belongs to the interior of $\mathcal{D}_\Lambda = \{\lambda : \Lambda(\lambda) < \infty\}$.

(2) Mixing properties.

(**A2**) $M_h = \int_0^1 (\alpha_X^{-1}(u))^h\, \mathrm{d}u < \infty$ for all $h > 0$.

(**A2**′) There exists $\gamma > 0$ and $\theta > 0$ such that $\alpha_{X,m} = O(\exp(-\gamma m^\theta))$.

(3) Approximation properties.

(**A3**) There exists a sequence $(S_n\Phi_n)_{n\geq 1}$ of functions from $E^{\mathbb{N}_+}$ to $\mathcal{Y}$ such that each $S_n\Phi_n$ depends only on the $2n$ first coordinates and

$$\delta_n = \sup_{z\in E^{\mathbb{N}_+}} \|S_n\Phi(z) - S_n\Phi_n(z)\|/n = o(1) \quad \text{as } n \to \infty.$$

Condition (**A3**) holds in particular if $S_n \Phi(X)$ is given by the Birkhoff sums of a function Φ defined on $E^{\mathbb{N}_+}$ and there exists a sequence of functions $(\Phi_n)_{n\geq 1}$ defined on $E^{\mathbb{N}_+}$ depending on the n first coordinates only, such that $\sup_{z\in E^{\mathbb{N}_+}} \|\Phi(z) - \Phi_n(z)\| = o(1)$ as n tends to ∞. If there exists an integer $p \geq 1$ such that Φ depends on the p first coordinates only, then one can take $\Phi_n = \Phi$ for n large enough, and then $\delta_n = 0$.

Now, we introduce the family of (random) probability measures for which we obtain large deviation results.

We fix an increasing sequence of positive integers $(k(n))_{n\geq 1}$, and for every $\omega \in \Omega$ and $n \geq 1$, we define

$$\mu_n^\omega = \frac{1}{k(n)} \sum_{j=1}^{k(n)} \delta_{x_{n,j}(\omega)} \quad \text{with } x_{n,j}(\omega) = S_n \Phi(T^{(j-1)n} X(\omega))/n.$$

In other words, for each Borel set $B \subset \mathbb{R}^d$, we have

$$\mu_n^\omega(B) = \frac{\#\{1 \leq j \leq k(n) : S_n \Phi(T^{(j-1)n} X(\omega))/n \in B\}}{k(n)}.$$

If $\mathcal{Y} = \mathbb{R}^d$, we also define

$$\Lambda_n^\omega(\lambda) = \frac{1}{n} \log \int_{\mathbb{R}^d} \exp(n\langle \lambda, x\rangle) \mathrm{d}\mu_n^\omega(x).$$

Now we start with results on the direct transfer of the LDP for $S_n \Phi(X)$ to the LDP for the local fluctuations of almost every realization of $S_n \Phi(X)$.

Theorem 2.1. *Assume* (**A1–3**).

(1) *Let $x \in \mathcal{D}_I$. If* $\liminf_{n\to\infty} \frac{\log k(n)}{n} > I(x)$ *then, with probability* 1,

$$\lim_{\epsilon\to 0^+} \liminf_{n\to\infty} \frac{1}{n} \log \mu_n^\omega(B(x,\epsilon)) = \lim_{\epsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \log \mu_n^\omega(B(x,\epsilon)) = -I(x).$$

If $\limsup_{n\to\infty} \frac{\log k(n)}{n} < I(x)$ *then there exists $\epsilon > 0$ such that, with probability* 1, *for n large enough the set $\{1 \leq j \leq k(n) : S_n \Phi(T^{(j-1)n} X(\omega))/n \in B(x,\epsilon)\}$ is empty.*

(2) *Let $x \in \mathcal{Y} \setminus \mathcal{D}_I$. With probability* 1,

$$\lim_{\epsilon\to 0^+} \limsup_{n\to\infty} \frac{1}{n} \log \mu_n^\omega(B(x,\epsilon)) = -I(x) = -\infty.$$

(3) *If $(\mu_n)_{n\geq 1}$ is exponentially tight, then so is $(\mu_n^\omega)_{n\geq 1}$ almost surely.*

Theorem 2.2. *Assume $(\mathcal{Y}, \|\ \|)$ is separable, as well as* (**A1–3**). *Suppose that* $\sup_{x\in\mathcal{D}_I} I(x) < \infty$ *and* $\liminf_{n\to\infty} \frac{\log k(n)}{n} > \sup_{x\in\mathcal{D}_I} I(x)$, *or that* $\lim_{n\to\infty} \frac{\log k(n)}{n} = \infty$. *With probability* 1, *$(\mu_n^\omega)_{n\geq 1}$ satisfies in $\mathcal{Y}$ the weak LDP with rate function I. If, moreover, $(\mu_n)_{n\geq 1}$ is exponentially tight, then $(\mu_n^\omega)_{n\geq 1}$ satisfies in $\mathcal{Y}$ the LDP with good rate function I.*

Next we give results concerning the transfer of convergence properties for Λ_n to convergence properties for Λ_n^ω. It is worth mentioning that under the assumptions of Theorem 2.2, if one has additional information like $\|S_n \Phi\|_\infty = O(n)$, then Varadhan's integral lemma (see [11] Th. 4.3.1) together with Theorem 2.2 directly provides the almost sure pointwise convergence of Λ_n^ω to Λ as $n \to \infty$.

Theorem 2.3. *Assume* (**A1′**) *and* (**A2–3**). *Let* $\boldsymbol{\lambda}_0 \in \mathring{\mathcal{D}}$ *at which* Λ *is differentiable and denote* $\nabla\Lambda(\lambda_0)$ *as* x_0.

(1) *If* $\liminf_{n\to\infty} \frac{\log k(n)}{n} > \Lambda^*(x_0)$ *then there exists* $r > 0$ *such that* $B(\lambda_0, r) \subset \mathcal{D}$ *and, with probability* 1, Λ_n^ω *converges uniformly to* Λ *over* $B(\lambda_0, r)$.

(2) *If* $\liminf_{n\to\infty} \frac{\log k(n)}{n} > \Lambda^*(x_0)$ *then, with probability* 1,

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty} \frac{1}{n}\log \mu_n^\omega(B(x_0, \epsilon)) = -\Lambda^*(x_0). \tag{2.5}$$

If $\limsup_{n\to\infty} \frac{\log k(n)}{n} < \Lambda^*(x_0)$, *there exists* $\epsilon > 0$ *such that, with probability* 1, *for n large enough the set* $\{1 \le j \le k(n) : S_n\Phi(T^{(j-1)n}X(\omega))/n \in B(x_0, \epsilon)\}$ *is empty.*

(3) *If* $\lim_{n\to\infty} \frac{\log k(n)}{n} = \Lambda^*(x_0)$ *and* $t \ge 0 \mapsto \Lambda(t\lambda_0)$ *is strictly convex at* 1 *then, with probability* 1, *for all* $t \ge 1$ *we have*

$$\lim_{n\to\infty} \Lambda_n^\omega(t\lambda_0) = \Lambda(\lambda_0) + (t-1)\langle\lambda_0, x_0\rangle. \tag{2.6}$$

Remark 2.1. (1) The almost sure large deviation equality (2.5) provided by Theorem 2.3(2) is a direct consequence of Theorem 2.3(1) and Theorem 1.5.

(2) In Theorem 1.1(3), since we consider a sequence of i.i.d. real valued random variables, if $t \ge 0 \mapsto \Lambda(t\lambda_0)$ is not strictly convex at 1 this means that X is constant, and the result obviously still holds.

The following result is a direct consequence of Theorem 2.3(1) and Theorem 1.4.

Corollary 2.1. *Assume* (**A1″**) *and* (**A2–3**).

Suppose that $\sup_{x\in\mathcal{D}_\Lambda} \Lambda^*(x) < \infty$ *and* $\liminf_{n\to\infty} \frac{\log k(n)}{n} > \sup_{x\in\mathcal{D}_\Lambda} \Lambda^*(x)$, *or* $\lim_{n\to\infty} \frac{\log k(n)}{n} = \infty$. *With probability 1,* Λ_n^ω *converges uniformly to* Λ *on the compact subsets of* $\mathcal{D}_\Lambda$, *hence the assertion of parts* (1) *and* (2) *of Theorem* 1.4 *hold for* $(\mu_n^\omega)_{n\ge 1}$. *If, moreover,* Λ *is essentially smooth and lower semi-continuous, the assertion of part* (3) *of Theorem* 1.4 *holds for* $(\mu_n^\omega)_{n\ge 1}$, *i.e.* $(\mu_n^\omega)_{n\ge 1}$ *satisfies in* $\mathbb{R}^d$ *the LDP with good rate function* $I = \Lambda^*$.

Next we want to measure more finely how big must be $k(n)$ for $\Lambda_n^\omega(\lambda)$ to converge to $\Lambda(\lambda)$ when Λ is smooth.

If $\mathcal{Y} = \mathbb{R}^d$, for any $n \ge 1$ and any subset B of $\mathcal{D}_\Lambda$ let

$$\begin{aligned} &\delta_n\Lambda(B) = \sup\{|\Lambda(\lambda) - \Lambda_n(\lambda)| : \lambda \in B\},\\ &\delta_n\Phi(B) = (\sup_{\lambda\in B}\|\lambda\|)\|S_n\Phi - S_n\Phi_n\|_\infty/n,\\ &\delta_n(\Lambda, \Phi)(B) = \delta_n\Lambda(B) + \delta_n\Phi(B), \end{aligned}$$

and if Λ is twice continuously differentiable, let

$$\begin{aligned} &\mathit{\Lambda}^*(B) = \sup\{\Lambda^*(\nabla\Lambda(\lambda)) : \lambda \in B\},\\ &\xi_1(B) = \sup\{\|\nabla\Lambda(\lambda)\| : \lambda \in B\},\\ &\xi_2(B) = \sup\left\{\frac{1}{2}\,{}^t\lambda \mathrm{D}^2\Lambda(\lambda)\lambda : \lambda \in B\right\}\\ &\xi(B) = \mathit{\Lambda}^*(B) + \xi_2(B), \end{aligned}$$

where $\mathrm{D}^2\Lambda(\lambda)$ stands for the Hessian matrix of Λ at λ.

If $B \subset \mathbb{R}^d$ and $\rho \in \mathbb{R}_+^*$ we define B_ρ as $\{\lambda \in \mathbb{R}^d : d(\lambda, B) \leq \rho\}$, where d stands for the Euclidean distance.

Theorem 2.4. *Suppose that* (**A1**$'$), (**A2**$'$) *and* (**A3**) *hold, and* Λ *is twice continuously differentiable over* $\mathcal{D}$. *Let* B *be a compact subset of* $\mathcal{D}$ *and let* $\rho > 0$ *such that* $B_\rho \subset \mathcal{D}$.

Suppose that there exists a positive sequence $(\epsilon_n)_{n\geq 1}$ *converging to* 0 *such that*

$$\sum_{n\geq 1} \exp(-\sqrt{\epsilon_n}[\log(k(n)) - n\Lambda^*(B)])\epsilon_n^{-(d+3/2)} \exp(3n[\xi(B_\rho)\epsilon_n + \delta_n(\Lambda, \Phi)(B_\rho)]) < \infty. \tag{2.7}$$

Let $\eta > 0$. *With probability* 1, *for* n *large enough,*

$$\max_{\lambda \in B} |\Lambda_n^\omega(\lambda) - \Lambda(\lambda)| \leq \mathcal{E}(n, \eta), \tag{2.8}$$

where $\mathcal{E}(n, \eta) = (\eta + 2\xi_1(B))\epsilon_n + \delta_n(\Lambda, \Phi)(B_\rho)$, *or* $\mathcal{E}(n, \eta) = \mathcal{E}(n) = \epsilon_n(1 + \epsilon_n)/n + \delta_n(\Lambda, \Phi)(B_\rho)$ *if* B *consists of only one point.*

Remark 2.2. (1) It follows easily from the proof of Theorem 2.4 (see (5.10)) that if $B = \{\lambda\}$, then in (2.7) one can replace $(d + 3/2)$ by $3/2$ to get the same conclusions as in Theorem 2.4.

(2) If the X_i are i.i.d, a simple modification of the proof using Lemma 5.1(2) rather than Lemma 5.1(1) makes it possible to replace $(d + 3/2)$ by $(d + 1)$ in (2.7).

(3) In the context described in Section 3.2, where X takes values in a symbolic space, $S_n\Phi(X)$ represents the Birkhoff sum of a continuous $\mathbb{R}^d$-valued potential Φ and the law of X is a Gibbs measure, we will give conditions under which both $\delta_n\Lambda(B_\rho)$ and $\delta_n\Phi(B_\rho)$ are $O(1/n)$. Then, a choice like $\epsilon_n = \gamma \log(n)/n$ and $\log(k(n))/n - \Lambda^*(B) \geq \sqrt{\gamma' \log(n)/n}$ with $\sqrt{\gamma\gamma'} > d + 5/2 + 3\gamma\xi(B_\rho)$ ensures that (2.7) holds and $\mathcal{E}(n, \eta) = O(\epsilon_n) = O(\log(n)/n)$.

Remark 2.3. As a first explicit example of situation to which Theorems 2.3 and 2.4 can be applied, let us consider products of random invertible matrices applied to a normalized vector. Let μ be a probability measure on $GL_m(\mathbb{R})$. Suppose that the support of μ generates a strongly irreducible and contracting semi-group (see Ch. III in [8] for the definition). Suppose also that $\exp(\tau \max(\log^+ \|x\|, \log^+ \|x^{-1}\|))$ is μ-integrable for some $\tau > 0$. Let $X = (X_i)_{i\geq 1}$ be a sequence of independent random matrices distributed according to μ. Fix a unit vector x and set $S_n\Phi(X) = \log \|X_n \cdots X_1 \cdot x\|$. There exists (see Ch V.6 in [8]) a neighborhood $\mathcal{D}$ of 0, independent of x, such that the limit Λ of Λ_n exists and is analytic on $\mathcal{D}$ (the derivative of Λ at 0 is the upper Lyapunov exponent associated with μ).

In the case where the X_i take values in $\mathbb{R}^d$ and are i.i.d, we also have the following improvement of Theorem 2.4.

Theorem 2.5. *Suppose that the* X_i *are i.i.d and take values in* $\mathbb{R}^d$. *Suppose also that* $S_n\Phi(X) = \sum_{k=1}^n X_i$, *and* $\Lambda(\lambda) = \log \mathbb{E}\exp(\langle\lambda, X_1\rangle)$ *is finite over a convex open subset of* $\mathbb{R}^d$.

Let B *be a compact subset of* $\mathcal{D}$ *and let* $\rho > 0$ *such that* $B_\rho \subset \mathcal{D}$. *Suppose that there exists a positive sequence* $(\epsilon_n)_{n\geq 1}$ *converging to* 0 *such that*

$$\sum_{n\geq 1} \exp(-\sqrt{\epsilon_n}[\log(k(n)) - n\Lambda^*(B)])\epsilon_n^{-(d+1)} \exp(\xi_2(B_\rho)\epsilon_n n) < \infty.$$

The same properties as in Theorem 2.4 hold, with $\mathcal{E}(n, \eta) = (\eta + 2\xi_1(B))\epsilon_n$, *or* $\mathcal{E}(n, \eta) = \mathcal{E}(n) = \epsilon_n(1 + \epsilon_n)/n$ *if* B *consists of only one point.*

Remark 2.4. If $B = \{\lambda\}$, in Theorem 2.5 we can take $\epsilon_n = \gamma \log(n)/n$ and $\log(k(n))/n - \Lambda^*(\nabla\Lambda(\lambda)) \geq \sqrt{\gamma' \log(n)/n}$ with $\sqrt{\gamma\gamma'} > d + 2 + \gamma \cdot \frac{1}{2}\, {}^t\lambda \mathrm{D}^2\Lambda(\lambda)\lambda$. Then $\mathcal{E}(n) \leq (1 + \epsilon_n)\epsilon_n/n = \gamma \log(n)/n^2 + \gamma^2 \log(n)^2/n^3$.

3. Examples

This section describes various contexts to which our results can be applied. We investigate applications to Brownian motion (Section 3.1), dynamical systems and number theory (Sections 3.2 and 3.3), branching random walks (Section 3.4) and Poissonian random walks (Section 3.5).

3.1. Fluctuations of the increments of Brownian motion

Let $(W_t)_{t\in[0,1]}$ be a d-dimensional standard Brownian motion defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let $(k(n))_{n\geq 1}$ be a sequence of positive integers. For each $n \geq 1$ and $1 \leq j \leq n$ we denote $[(j-1)/n, j/n]$ by $J_{n,j}$ and the increment of W over the interval $J_{n,j}$ is then denoted by $\Delta W(J_{n,j})$.

For every $\omega \in \Omega$ and $n \geq 1$, define

$$\mu_n^\omega = \frac{1}{k(n)} \sum_{j=1}^{k(n)} \delta_{x_{n,j}(\omega)}, \quad \text{with } x_{n,j}(\omega) = (k(n)/n)^{1/2} \Delta W(J_{k(n),j}).$$

In other words, for each Borel set $B \subset \mathbb{R}^d$, we have

$$\mu_n^\omega(B) = \frac{\#\{1 \leq j \leq k(n) : (k(n)/n)^{1/2} \Delta W(J_{k(n),j}) \in B\}}{k(n)}.$$

The following result is essentially a refinement of Theorem 2.3 applied to a sequence of independent centered Gaussian vectors with covariance matrix the identity. We will give a short proof in Section 6.1.

Theorem 3.1. *Let $R > 0$. Suppose that there exists a positive sequence $(\epsilon_n)_{n\geq 1}$ converging to 0 such that*

$$\sum_{n\geq 1} \exp(-\sqrt{\epsilon_n}[\log(k(n)) - n(1 + \sqrt{\epsilon_n}) R^2/2]) \epsilon_n^{-1} < \infty. \tag{3.1}$$

With probability 1, for every Borel subset Γ of $\mathring{B}(0, R)$, (1.6) holds for $(\mu_n^\omega)_{n\geq 1}$, with rate function $I(x) = \|x\|^2/2$.

The choice $\epsilon_n = \gamma \log(n)/n$ and $\log(k(n))/n - (1+\sqrt{\epsilon_n})R^2/2 \geq \sqrt{\gamma' \log(n)/n}$ with $\sqrt{\gamma\gamma'} > 2$ yields (3.1).

We also have a functional result based on the LDP established by Schilder (see [11] Th. 5.2.3): for $n \geq 1$, let ν_n stand for the distribution of $W/\sqrt{n}$ as a random element of $C_0([0, 1])$, the space of $\mathbb{R}^d$-valued continuous functions ϕ over $[0, 1]$ such that $\phi(0) = 0$. Then $(\nu_n)_{n\geq 1}$ satisfies in $C_0([0, 1])$ the LDP with good rate function

$$I(\phi) = \begin{cases} \dfrac{1}{2} \displaystyle\int_0^1 \phi'(t)^2 \, \mathrm{d}t & \text{if } \phi \in H^1 \\ \infty & \text{otherwise,} \end{cases}$$

where H^1 stands for Sobolev space of absolutely continuous elements of $C_0([0,1])$ with square integrable derivative.

It follows from Schilder's theorem that if $X = (X_i)_{i\geq 1}$ is a sequence of independent standard Brownian motions and $S_n = X_1 + \cdots + X_n$, the distributions of the variables S_n/n also satisfy in $C_0([0,1])$ the LDP with rate I. Consequently we get almost surely the LDP with rate I for the local fluctuations of S_n in the sense of Theorem 2.2. This essentially yields the following result.

For each $n \geq 1$ and $1 \leq j \leq k(n)$ denote by $W_{k(n),j}$ the standard Brownian motion $t \in [0,1] \mapsto k(n)^{1/2}(W((t+(j-1))/k(n)) - W((j-1)/k(n)))$. For every $\omega \in \Omega$ and $n \geq 1$, define

$$\mu_n^\omega = \frac{1}{k(n)} \sum_{j=1}^{k(n)} \delta_{x_{k(n),j}(\omega)}, \quad \text{with } x_{k(n),j}(\omega) = \frac{W_{k(n),j}}{n^{1/2}}.$$

Theorem 3.2. *Suppose that* $\lim_{n\to\infty} \frac{\log k(n)}{n} = \infty$. *With probability* 1, $(\mu_n^\omega)_{n\geq 1}$ *satisfies in* $C_0([0,1])$ *the LDP with good rate function* I.

Remark 3.1. It is possible to combine the ideas developed in this paper with those of [30] to obtain results in the spirit of Theorem 3.2 for some Lévy processes with jumps.

3.2. Local fluctuations of Birkhoff sums and products of matrices with respect to Gibbs measures

Let Σ_m stand for the one sided symbolic space over a finite alphabet of cardinality $m \geq 2$: $\Sigma_m = \{0, \ldots, m-1\}^{\mathbb{N}_+}$. The set Σ_m is endowed with the shift operation $T(\{t_n\}_{n=1}^\infty) = \{t_{n+1}\}_{n=1}^\infty$. Let A be a $m \times m$ matrix with all entries equal to 0 and 1 and such that A^p is positive for some $p \geq 1$. Then let (Σ_A, T) be the associated topologically mixing subshift of finite type of (Σ_m, T), i.e. $\Sigma_A = \{t \in \Sigma_m : \forall n \geq 1,\ A_{t_n,t_{n+1}} = 1\}$.

We denote by $\mathcal{M}(\Sigma_A, T)$ the set of invariant probability measures under T.

For $n \geq 1$ we define $\Sigma_{A,n} = \{(t_1 \ldots t_n) \in \{0, \ldots, m-1\}^n : \forall\, 1 \leq k \leq n-1,\ A_{t_k,t_{k+1}} = 1\}$.

If $t \in \Sigma_A$ and $n \geq 1$ we denote $t_1 \cdots t_n$ by $t_{|n}$ and for $w \in \Sigma_{A,n}$ the cylinder $\{t \in \Sigma_A : t_{|n} = w\}$ is denoted $[w]$.

The set Σ_A is also endowed with the standard ultra-metric distance $d(t,s) = m^{-|t\wedge s|}$, where $|t \wedge s| = \sup\{n : t_{|n} = s_{|n}\}$.

If ψ is a continuous function from Σ_A to $\mathbb{R}$, the topological pressure of ψ is defined as $P(T,\psi) = \sup\{\nu(\psi) + h_\nu(T) : \nu \in \mathcal{M}(\Sigma_A, T)\}$, and one has (see [9])

$$P(T,\psi) = \lim_{n\to\infty} \frac{1}{n} \sum_{w\in\Sigma_{A,n}} \sup_{y\in[w]} \exp(S_n\psi(y)).$$

We say that ψ satisfies the bounded distortion property if

$$\sup_{n\geq 1} v_n < \infty, \text{ where } v_n = \sup_{\substack{t,s\in\Sigma_A \\ t_{|n}=s_{|n}}} |S_n\psi(t) - S_n\psi(s)| < \infty.$$

In this case, it is well known that $\sup\{\nu(\psi) + h_\nu(T) : \nu \in \mathcal{M}(\Sigma_A, T)\}$ is attained at a unique and ergodic measure called the equilibrium state of ψ (see [9,35]). We will denote it by ν_ψ. This measure is a Gibbs measure, in the sense that there exists a constant $C > 0$ such that

$$\begin{aligned}&\forall n \geq 1,\ \forall t \in \Sigma_A,\ C^{-1}\exp(S_n\psi(t) - nP(T,\psi)) \leq \nu_\psi([t_{|n}]) \\ &\quad \leq C\exp(S_n\psi(t) - nP(T,\psi)).\end{aligned} \tag{3.2}$$

Moreover, if Φ is a continuous mapping from Σ_A to $\mathbb{R}^d$ such that each component of Φ satisfies the bounded distortion property, then $\lambda \in \mathbb{R}^d \mapsto P(T, \langle\lambda, \Phi\rangle)$ is a C^1 mapping from $\mathbb{R}^d$ to $\mathbb{R}$ (see [36,5]).

3.2.1. Results for Birkhoff sums

We fix a real valued potential ψ on Σ_A satisfying the bounded distortion property. Then, the process X defined as the identity map of Σ_A is stationary with respect to the ergodic measure ν_ψ. We also fix Φ, a continuous mapping from Σ_A to $\mathbb{R}^d$ and define $(S_n\Phi(X))_{n\geq 1}$ as the sequence of Birkhoff sums of Φ.

Thus, setting $(\Omega, \mathbb{P}) = (\Sigma_A, \nu_\psi)$, the quantities introduced in Section 2 take the following form. For all $n \geq 1$, $B \in B_{\mathbb{R}^d}$ and $\lambda \in \mathbb{R}^d$,

$$\mu_n(B) = \nu_\psi(\{t \in \Sigma_A : S_n\Phi(t)/n \in B\})$$

and

$$\Lambda_n(\lambda) = \frac{1}{n}\log \mathbb{E}\exp(\langle\lambda, S_n\Phi(X)\rangle) = \frac{1}{n}\int_{\Sigma_A} \exp(\langle\lambda, S_n\Phi(t)\rangle)\,\mathrm{d}\nu_\psi(t).$$

Also, μ_n^ω and Λ_n^ω are denoted μ_n^t and Λ_n^t respectively and we have for $t \in \Sigma_A$, $n \geq 1$, and $\lambda \in \mathbb{R}^d$

$$\mu_n^t(B) = \frac{\#\{1 \leq j \leq k(n) : S_n\Phi(T^{(j-1)n}t)/n \in B\}}{k(n)}$$

and

$$\Lambda_n^t(\lambda) = \frac{1}{n}\log\int_{\mathbb{R}^d} \exp(n\langle\lambda, x\rangle)\mathrm{d}\mu_n^t(x).$$

Due to the Gibbs properties of ν_ψ (3.2), $\Lambda(\lambda) = \lim_{n\to\infty}\Lambda_n(\lambda)$ exists and takes the form

$$\Lambda(\lambda) = P(T, \psi + \langle\lambda, \Phi\rangle) - P(T, \psi).$$

If, moreover, each component of Φ satisfies the bounded distortion property then Λ is C^1. Thus, condition (**A1**$''$) (hence (**A1**$'$)) hold with $\mathcal{D}_\Lambda = \mathbb{R}^d$. Moreover, $\delta_n\Lambda(B) = O(1/n)$ for bounded sets B.

For (**A2**) to hold we must ask some mixing properties of ν_ψ. It is quite simple to see that (**A2**) holds under the stronger assumption that there exists $\gamma > 0$ and $\theta > 1$ such that $\alpha_{X,m} = O(\exp(-\gamma\log(m)^\theta))$. Then, due to Theorem 1.11 in [3], (**A2**) holds as soon as the modulus of continuity of ψ, namely $\kappa(\psi, \cdot)$ satisfies $\kappa(\psi, \delta) = O(\exp(-\gamma(\log|\log(\delta)|)\,|^\theta))$ as $\delta \to 0$ for some $\gamma > 0$ and $\theta > 1$. Also (**A2**$'$) holds as soon as $\kappa(\psi, \delta) = O(\exp(-\gamma|\log(\delta)|^\theta))$ as $\delta \to 0$ for some $\gamma > 0$ and $\theta > 0$.

The function Φ being continuous on the compact set (Σ_A, d), (**A3**) always holds since we can always approximate Φ by a function Φ_n depending only on $(t_1, \ldots, t_n)$ so that $\|S_n\Phi - S_n\Phi_n\|_\infty \leq \sum_{k=1}^n \kappa(\Phi, m^{-k}) = o(n)$.

Thus under the above conditions on Φ and ψ assuring (**A1**$''$) and (**A2**) Theorem 2.3 and Corollary 2.1 can be applied to this context and provide information regarding the convergence of Λ_n^t to Λ for ν_ψ-almost every t. If, moreover, we assume that ψ and the components of Φ are Hölder continuous, then Λ is analytic (see for instance Th. 5 in [35]) and (**A2**$'$) holds, so that we can apply Theorem 2.4.

In fact, even if Φ is only supposed continuous, $(\mu_n)_{n\geq 1}$ satisfies in $\mathbb{R}^d$ the LDP with good rate function

$$I(x) = \begin{cases} \inf\{P(T,\psi) - (h_\nu(T) + \nu(\psi)) : \nu \in \mathcal{M}(\Sigma_A, T), \\ \qquad \nu(\Phi) = x\} = \Lambda^*(x) & \text{if } x \in \mathcal{D}_I \\ \infty & \text{otherwise,} \end{cases} \tag{3.3}$$

where $\mathcal{D}_I = \{\nu(\Phi) : \nu \in \mathcal{M}(\Sigma_A, T)\}$, and I is bounded over the compact convex set $\mathcal{D}_I$. This LDP essentially follows from Theorem 6 of [44] (which deals with Hölder potentials), and the duality between the pressure and entropy functions (see [15,16,41,20] for details and related works). Thus (**A1**) holds. It follows that we can apply Theorem 2.2 and transfer the previous LDP to the local fluctuations of $S_n\Phi$:

Theorem 3.3. *If* $\liminf_{n\to\infty} \frac{\log k(n)}{n} > \sup_{x\in\mathcal{D}_I} \Lambda^*(x)$*, then for* ν_ψ*-almost every* t*, the sequence* $(\mu_n^t)_{n\geq 1}$ *satisfies in* $\mathbb{R}^d$ *the LDP with good rate function given by* (3.3)*.*

Thus, we can also deal with the cases where the function Λ is non-differentiable at some $\lambda \in \mathbb{R}^d$ because $\langle \lambda, \Phi\rangle + \psi$ have at least two equilibrium states with distinct entropies (see [35] p. 52 for instance).

Some geometric applications. The previous results have applications to geometric realizations of (Σ_A, T), for instance on repellers of topologically mixing $C^{1+\epsilon}$ conformal maps of Riemannian manifolds. For such a repeller (J, f), they make it possible to describe the local fluctuations of $S_n \log\|Df\|$ almost everywhere with respect to any enough mixing Gibbs measure on (J, f); this means that while with respect to such a measure ν_ψ one observes on almost every orbit an expansion ruled by a fixed Lyapunov exponent equal to $\nu_\psi(\log\|Df(x)\|)$, we can finely quantify local fluctuations with respect to this global property. The same can be done along the stable and unstable manifolds on locally maximal invariant sets of topologically mixing Axiom A diffeomorphisms (see [9,21] for details on these dynamical systems).

Another application concerns the harmonic measure on planar Cantor repellers of $C^{1+\epsilon}$ conformal maps f; recall that given such a repeller J, this measure is the probability measure μ such that for each $t \in J$ and $r > 0$, $\mu(B(t,r))$ is the probability that a planar Brownian motion started at ∞ attains J for the first time at a point of $B(t,r)$. It turns out that μ is equivalent to the equilibrium state μ_φ of a Hölder potential φ on J (see [10] or [27]). Given another enough mixing Gibbs measure ν_ψ, the ergodic theorem ensures that $\lim_{r\to 0^+} \log\mu_\varphi(B(t,r))/\log(r) = (P(\varphi) - \nu_\psi(\varphi))/\nu_\psi(\log\|Df\|)$ for ν_ψ-almost every t. Then, our result yields information on the fluctuations with respect to this behavior. Indeed, one can use the coding of (J, f) by a subshift of finite type thanks to a Markov partition and apply our results to the pair $(S_n(\varphi - P(\varphi)), S_n \log\|Df\|)$. If we remember the origin of φ, this yields information on the local fluctuations of the Brownian motion around ν_ψ-almost every t. This can be made more explicit in the case that J is self-similar and homogeneous, for instance when $J = K^2$ with K the middle third Cantor set. There our results provide, for ν_ψ-almost every t, information on the distributions of the values

$$\frac{\log\mu_\varphi(C_{jn}(t))}{\log\mu_\varphi(C_{(j-1)n}(t))} = 1 + \frac{S_n(\varphi - P(\varphi))(f^{(j-1)n}(t))}{jn\nu_\psi(\varphi - P(\varphi))} + O\left(\frac{1}{jn}\right), \quad 1 \leq j \leq k(n),$$

where $C_k(t)$ is the triadic cube of generation k containing t.

The previous interpretations of our results about the local behavior of Gibbs measures can be extended to the case of Axiom A diffeomorphisms invoked above.

Thus, to summarize, while [31,32,44] provide large deviations with respect to the almost sure asymptotic behavior of Birkhoff sums on a hyperbolic invariant set endowed with a Gibbs measure, our results provide a natural complement by describing the fluctuations with respect to this behavior on almost every orbit viewed by this measure.

The next two subsections briefly discuss extensions to norms of Birkhoff products of matrices of the previous properties of Birkhoff sums of potentials.

3.2.2. Birkhoff products of positive matrices

Suppose that M is a mapping from Σ_A to the set of positive square matrices of order $d \geq 1$, and fix an enough mixing Gibbs measure ν_ψ. If the components $M_{i,j}$ are so that $\log(M_{i,j})$ has the bounded distortion property then, one can apply Theorem 2.3 to $S_n\Phi(X(t)) = S_n\Phi(t) = \log\|M(t)M(T(t))\cdots M(T^{n-1}(t))\|$ with respect to ν_ψ. Indeed, the convergence of $\Lambda_n(\lambda)$ for $\lambda \in \mathbb{R}$ comes from the Gibbs property of ν_ψ and the subadditivity and superadditivity properties of $S_n\Phi(X)$, and the differentiability of Λ comes from the variational principle for subadditive potentials (see [18] for instance). If the components of M are only supposed continuous $\Lambda_n(\cdot)$ still converges but may be non-differentiable. It is then possible to extend the result explained in the previous subsection and show that (**A1**) holds for $S_n\Phi$ with the good rate function I still satisfying (3.3) and bounded over $\mathcal{D}_I$. The only difference is that here $\nu(\Phi)$ is defined as $\lim_{n\to\infty} n^{-1}\int_{\Sigma_A} S_n\Phi(t)\,\mathrm{d}t$.

3.2.3. Bernoulli products of invertible matrices

Suppose that we are given $M_1,\ldots,M_m$, m matrices of $GL(d,\mathbb{C})$ such that there is no proper non-zero linear subspace V of $\mathbb{C}^d$ such that $M_i(V) \subset V$. Then, define $M(t) = M_{t_1}$ and $S_n\Phi(X(t)) = S_n\Phi(t) = \log\|M_{t_1}\cdots M_{t_n}\|$ for $t \in \Sigma_m$. Nice superadditivity and subadditivity properties (see [19]) make it possible to extend the results of the previous section to this context. We do not enter into the details.

3.3. Local fluctuations in the continued fraction expansion of Lebesgue-almost every point

The interval $[0, 1)$ is endowed with the dynamics of the Gauss transformation $f(0) = 0$, $f(t) = 1/t - \lfloor 1/t \rfloor$ if $t \in (0, 1)$. Then, the continued fraction expansion of an irrational number $t \in (0, 1)$ is represented by the sequence $[a_1(t); a_2(t); \ldots; a_n(t); \ldots]$, where $a_1(t) = \lfloor 1/t \rfloor$ and $a_n(t) = a_1(f^{n-1}(t)) = \lfloor 1/f^{n-1}(t) \rfloor$. The Gauss measure μ_G whose density with respect to the Lebesgue measure on $[0, 1)$ is $1/(1+t)\log(2)$ is ergodic with respect to f, and it possesses the strong mixing properties required in (**A2**) (see [7] for instance). Now let $\Phi(t) = \log a_1(t)$ for $t \in (0, 1)$. An application of the Birkhoff ergodic theorem proves that for Lebesgue-almost every t, one has $S_n\Phi(t) = \sum_{k=0}^{n-1}\log a_k(t) \sim n\int_0^1 \log a_1(t)\,\mathrm{d}\mu_G(t)$.

Here we are concerned with the limit of $\Lambda_n(\lambda) = n^{-1}\log\int_0^1 \exp(\lambda S_n\Phi(t))\,\mathrm{d}\mu_G(t)$ whenever it exists. For each $n \geq 1$ and each sequence $a_1,\ldots,a_n$ of integers let us denote by $I_{a_1,\ldots,a_n}$ the interval $\{t \in [0, 1) : [a_1(t); a_2(t); \ldots; a_n(t)] = [a_1; a_2; \ldots; a_n]\}$. It is clear that the question reduces to studying $n^{-1}\log\sum_{(a_1,\ldots,a_n)\in(\mathbb{N}_+)^n}(a_1\cdots a_n)^\lambda |I_{a_1,\ldots,a_n}|$; this sequence converges for $\lambda < 1$ to a limit $\Lambda(\lambda)$ analytic in λ (see Section 4 of [17]). Consequently, Theorem 2.3, Corollary 2.1 and Theorem 2.4 provide large deviation properties for the local fluctuations of $\log(a_1(t)) + \cdots + \log(a_n(t))$ almost everywhere with respect to the Lebesgue measure.

The previous example can be generalized by studying the local fluctuations of the Birkhoff sums associated with good potentials on the symbolic space over an infinite alphabet with respect to enough mixing Gibbs measures. We refer the reader to [17] and [37] for further examples and references.

3.4. Local fluctuations of branching random walks (BRW) with respect to generalized branching measures

Let $(N, (\psi_1, \Phi_1), (\psi_2, \Phi_2), \ldots)$ be a random vector taking values in $\mathbb{N}_+ \times (\mathbb{R} \times \mathbb{R}^d)^{\mathbb{N}_+}$. In the sequel, the distribution of N will define a supercritical Galton–Watson tree, on the boundary of which will live a Mandelbrot measure determined by $(\psi_1, \psi_2, \ldots)$, with respect to which we will look almost everywhere at the local fluctuations of a branching random walk whose distribution is determined by $(\Phi_1, \Phi_2, \ldots)$. Here, $(\psi_1, \psi_2, \ldots)$ and $(\Phi_1, \Phi_2, \ldots)$ play roles analogous to the potentials ψ and Φ in the previous section.

Let $\{(N_{u0}, (\psi_{u1}, \Phi_{u1}), (\psi_{u2}, \Phi_{u2}), \ldots)\}_u$ be a family of independent copies of the vector $(N, (\psi_1, \Phi_1), (\psi_2, \Phi_2), \ldots)$ indexed by the finite sequences $u = u_1 \cdots u_n$, $n \geq 0$, $u_i \in \mathbb{N}_+$ ($n = 0$ corresponds to the empty sequence denoted $\emptyset$), and let T be the Galton–Watson tree with defining elements $\{N_u\}$: we have $\emptyset \in \mathsf{T}$ and, if $u \in \mathsf{T}$ and $i \in \mathbb{N}_+$ then ui, the concatenation of u and i, belongs to T if and only if $1 \leq i \leq N_u$. Similarly, for each $u \in \bigcup_{n\geq 0} \mathbb{N}_+^n$, denote by $\mathsf{T}(u)$ the Galton–Watson tree rooted at u and defined by the $\{N_{uv}\}$, $v \in \bigcup_{n\geq 0} \mathbb{N}_+^n$.

The probability space over which these random variables are built is denoted $(\Upsilon, \mathcal{A}, \mathsf{P})$, and the expectation with respect to P is denoted E.

Let us define the $\mathbb{R} \cup \{\infty\}$-valued convex mapping

$$\mathsf{L} : \lambda \in \mathbb{R}^d \mapsto \log \mathsf{E}\left(\sum_{i=1}^{N} \exp(\psi_i + \langle \lambda, \Phi_i \rangle)\right).$$

We assume that

$$\mathsf{E}\left(\sum_{i=1}^{N} \exp(\psi_i)\right) = 1, \qquad \log \mathsf{E}\left(\sum_{i=1}^{N} \psi_i \exp(\psi_i)\right) < 0 \quad \text{and}$$
$$\mathsf{E}\left(\left(\sum_{i=1}^{N} \exp(\psi_i)\right) \log^+ \left(\sum_{i=1}^{N} \exp(\psi_i)\right)\right) < \infty.$$

Then, it is known (see [28,23,26]) that for each $u \in \bigcup_{n\geq 0} \mathbb{N}_+^n$, the sequence

$$Y_n(u) = \sum_{v=v_1\cdots v_n \in \mathsf{T}(u)} \exp(\psi_{uv_1} + \cdots + \psi_{uv_1\cdots v_n})$$

is a positive uniformly integrable martingale of expectation 1 with respect to the natural filtration. We denote by $Y(u)$ its P-almost sure limit. By construction, the random variables so obtained are identically distributed and positive. Also, the Galton–Watson tree T is supercritical.

Now, for each $u \in \bigcup_{n\geq 0} \mathbb{N}_+^n$, we denote by $[u]$ the cylinder $u \cdot {\mathbb{N}_+}^{\mathbb{N}_+}$ and define

$$\nu([u]) = \mathbf{1}_T(u) \exp(\psi_{u_1} + \cdots + \psi_{u_1\cdots u_n})\, Y(u).$$

Due to the branching property $Y(u) = \sum_{i=1}^{N(u)} \exp(\psi_{ui}) Y(ui)$, this yields a non-negative additive function of the cylinders, so it can be extended into a random measure ν_γ ($\gamma \in \Upsilon$) on $\mathbb{N}_+^{\mathbb{N}_+}$ endowed with the Borel σ-field $\mathcal{B} = \mathcal{B}(\mathbb{N}_+^{\mathbb{N}_+})$. This measure has $\partial T = \bigcap_{n\geq 0} \bigcup_{u=u_1\cdots u_n \in T} [u]$ as support.

Now, let $\Omega = \Upsilon \times \mathbb{N}_+^{\mathbb{N}_+}$. We can define on $(\Omega, \mathcal{A} \otimes \mathcal{B})$ the probability measure

$$\mathbb{P}(A) = \int_\Upsilon \int_{{\mathbb{N}_+}^{\mathbb{N}_+}} \mathbf{1}_A(\gamma, t) \mathrm{d}\nu_\gamma(t) \mathrm{d}\mathsf{P}(z).$$

Then, it is known (see [24] for instance) that the random variables $X_n(\gamma, t) = \Phi_{t_1\cdots t_n}(\gamma)$ are i.i.d. with respect to $\mathbb{P}$. If, moreover, $\nabla \mathsf{L}(0)$ exists then it equals $\mathbb{E}(X_1)$ and $(X_1 + \cdots X_n)/n$ tends to $\nabla \mathsf{L}(0)$ $\mathbb{P}$-almost surely. In terms of the BRW $\sum_{i=1}^n \Phi_{t_1\cdots t_i}$ on T, this means that with P-probability 1, for ν_γ-almost every $t \in \partial T$, we have $\lim_{n\to\infty} \sum_{i=1}^n \Phi_{t_1\cdots t_i}(\gamma)/n = \nabla \mathsf{L}(0)$.

Moreover, in the present context, if we set $X = (X_i)_{i\ge 1}$, since the X_i are i.i.d. we have $\Lambda(\lambda) = \log \mathbb{E} \exp(\langle \lambda, X_1\rangle) = \mathsf{L}(\lambda)$. Consequently, if L is finite on an open convex subset $\mathcal{D}$ of $\mathbb{R}^d$, local fluctuations of the BRW $\sum_{i=1}^n \Phi_{t_1\cdots t_i}$ are described P-almost surely ν_γ-almost everywhere thanks to Theorem 2.3, Corollary 2.1 and Theorem 2.4. When $\Phi_i \in \{0, 1\}$ for all $i \ge 1$, this is related to percolation on the Galton–Watson tree T (see [25]).

3.5. *Local fluctuations of Poissonian random walks and covering numbers with respect to compound Poisson cascades*

As in the previous section, the probability space over which we are going to define random variables is denoted $(\Upsilon, \mathcal{A}, \mathsf{P})$, and the expectation with respect to P is denoted E.

Let $\xi > 0$ and $\mathcal{P}$ a Poisson point process in $\mathbb{R} \times (0, 1]$ with intensity Λ given by

$$\Lambda(\mathrm{d}s\mathrm{d}\lambda) = \frac{\xi \mathrm{d}s\mathrm{d}\lambda}{\lambda^2}.$$

For every $(s, \lambda) \in \mathcal{P}$ let $J(s, \lambda) = (s, s + \lambda)$. The question of knowing whether $\mathbb{R}_+ \setminus \{0\}$ is or not almost surely covered by the intervals $J(s, \lambda)$ has been raised in [29] in connexion with a similar problem previously raised in [12] for random arcs on the circle. These problems have been solved in [38,39] (see also [22] for further information on this question). Then, works [14,4] have been dedicated to the geometric heterogeneity of the asymptotic behavior of the covering numbers defined as follows (in fact, all the works mentioned above consider more generally the case of Poisson intensities invariant by horizontal translation). Here we rather look at local fluctuations of these numbers.

For every $t \in [0, 1]$ and $n \ge 0$, the covering number of t at height e^{-n} by the Poissonian intervals $J(s, \lambda)$ is defined as

$$N_n(\gamma, t) = \sum_{(s,\lambda)\in\mathcal{P},\ \lambda > \mathrm{e}^{-n}} \mathbf{1}_{\{J(s,\lambda)\}}(t) = \#\big\{(s, \lambda) \in \mathcal{P} :\ \lambda > \mathrm{e}^{-n},\ t \in J(s, \lambda)\big\} \quad (\gamma \in \Upsilon).$$

For every $t \in [0, 1]$, this covering number can be seen as the “Poissonian” random walk $S_n(\gamma, t) = X_1(\gamma, t) + \cdots + X_n(\gamma, t)$ associated with the random variables $X_i(\gamma, t)$ defined as

$$\begin{aligned} X_i(\gamma, t) &= \sum_{\substack{(s,\lambda)\in\mathcal{P},\\ \mathrm{e}^{-i} < \lambda \le \mathrm{e}^{-(i-1)}}} \mathbf{1}_{\{J(s,\lambda)\}}(t) \\ &= \#\{(s, \lambda) \in \mathcal{P} :\ \mathrm{e}^{-i} < \lambda \le \mathrm{e}^{-(i-1)},\ t \in J(s, \lambda)\big\} \quad (i \ge 1). \end{aligned}$$

The choice of Λ ensures that the $X_i(\cdot, t)$ are i.i.d. We can describe the fluctuations of $S_n(\gamma, t)$ thanks to Theorems 2.3 and 2.5 by considering random measures on $\mathbb{R}_+$, namely compound Poisson cascades [6]. In fact, the invariance by horizontal translation of the constructions makes it possible to restrict ourselves to [0, 1] without loss of generality.

It turns out that we can also describe a more general model of Poissonian random walks in the spirit of branching random walks. To this, we consider a random vector $(\psi, \Phi) \in \mathbb{R} \times \mathbb{R}^d$, and to each $(s, \lambda) \in \mathcal{P}$ we associated a copy $(\psi_{(s,\lambda)}, \Phi_{(s,\lambda)})$ of (ψ, Φ) in such a way that these random variables are independent and independent of $\mathcal{P}$.

For each $t \in [0, 1]$ and $\phi \in \{\psi, \Phi\}$ we consider the random variables

$$X_i^{\phi}(\gamma, t) = \sum_{\substack{(s,\lambda)\in\mathcal{P},\\ e^{-i}<\lambda\le e^{-(i-1)}}} \phi_{(s,\lambda)} \quad (i \ge 1,\ \gamma \in \Upsilon)$$

as well as the Poissonian random walk $S_n^{\phi}(\gamma, t) = X_1^{\phi}(\gamma, t)+\cdots+X_n^{\phi}(\gamma, t)$. An easy calculation shows that for any $(q, \lambda) \in \mathbb{R}\times\mathbb{R}^d$, for every $t \in \mathbb{R}_+$ one has $\mathsf{E}\exp(qS_n^{\psi}(\cdot, t)+\langle\lambda, S_n^{\Phi}(\cdot, t)\rangle) = \exp(n\xi\mathsf{E}\big(\exp(q\psi + \langle\lambda, \Phi\rangle) - 1)\big)$.

We define over $[0, 1]$ the sequence of random measures introduced in [6] as

$$\begin{aligned}\nu_{\gamma,n}(\mathrm{d}t) &= (\mathsf{E}\exp(S_n^{\psi}(\cdot, t)))^{-1}\exp(S_n^{\psi}(\gamma, t))\,\mathrm{d}t\\ &= \exp(S_n^{\psi}(\gamma, t) - n\xi\mathsf{E}(\exp(\psi) - 1))\mathrm{d}t.\end{aligned} \tag{3.4}$$

Let $\tau(q) = (1-\xi)(1-q) + \xi\mathsf{E}\big(\exp(q\psi) - q\exp(\psi)\big)$. We assume that $\tau'(1) < 0$. Then, for P-almost every γ, $\nu_{\gamma,n}$ converges in the weak-star topology to a fully supported measure ν_γ over $[0, 1]$, and whose total mass has expectation 1 (see [6]). We can defined on $\Omega = \Upsilon \times [0, 1]$ endowed with $\mathcal{A}\otimes\mathcal{B}([0, 1])$ the probability measure

$$\mathbb{P}(A) = \int_{\Upsilon}\int_{[0,1]} \mathbf{1}_A(\gamma, t)\mathrm{d}\nu_\gamma(t)\mathrm{d}\mathsf{P}(\gamma).$$

Let

$$L : \lambda \in \mathbb{R}^d \mapsto \xi\mathsf{E}(\exp(\psi + \langle\lambda, \Phi\rangle) - 1).$$

The random variables $X_i^{\phi}(\gamma, t)$ are i.i.d with respect to $\mathbb{P}$, and it is not difficult to see that $\Lambda(\lambda) = \mathsf{L}(\lambda)$. Thus, if L is finite on an open convex subset $\mathcal{D}$ of $\mathbb{R}^d$, Theorem 2.3, Corollary 2.1 and Theorem 2.5 applied to $X = (X_i)_{i\ge1}$ with respect to $\mathbb{P}$ provide a description of the local fluctuations of $S_n^{\Phi}(\gamma, t)$, P-almost surely, for ν_γ-almost every t.

4. Conjecture on the “randomness” of fundamental constants

As mentioned in the introduction, our results lead us to formulate a new conjecture regarding how in any integer basis m the digits of fundamental constants such as the number Pi or the Euler constant look like almost every realization of a sequence of i.i.d random variables uniformly distributed in $\{0, \ldots, m-1\}$. This conjecture implies the normality property.

Recall the notations of Section 3.2. Consider a $\mathbb{R}^d$-valued continuous potential Φ defined on $\Sigma_m = \{0, \ldots, m-1\}^{\mathbb{N}_+}$ endowed with the shift operation denoted T. Consider a sequence $(k(n))_{n\ge1}$ of positive integers. Recall that in Section 3.2.1 we have defined for $t \in \Sigma_m$ the sequence of Borel measures $(\mu_n^t)_{n\ge1}$ and logarithmic generating functions $(\Lambda_n^t)_{n\ge1}$ as

$$\mu_n^t = \frac{1}{k(n)}\sum_{j=1}^{k(n)} \delta_{x_{n,j}(t)} \quad \text{with } x_{n,j}(t) = S_n\Phi(T^{(j-1)n}t)/n$$

and

$$\Lambda_n^t(\lambda) = \frac{1}{n}\log\int_{\mathbb{R}^d} \exp(n\langle\lambda, x\rangle)\mathrm{d}\mu_n^t(x). \tag{4.1}$$

Consider now the potential $\psi = 0$ and the associated equilibrium state ν_ψ, i.e. the measure of maximal entropy on (Σ_m, T). We have $P(\psi) = \log(m)$.

The process $X = (X_i)_{i\geq 1}$ defined on the probability space (Σ_m, ν_ψ) as $X(t) = (t_i)_{i\geq 1}$ is a sequence of i.i.d random variables uniformly distributed in $\{0, \dots, m-1\}$, and the rate function I provided by (3.3) takes the form

$$I(x) = \begin{cases} \inf\{\log(m) - h_\nu(T) : \nu \in \mathcal{M}(\Sigma_A, T),\ \nu(\Phi) = x\} = \Lambda^*(x) & \text{if } x \in \mathcal{D}_I \\ \infty & \text{otherwise,} \end{cases}$$

where $\mathcal{D}_I = \{\nu(\Phi) : \nu \in \mathcal{M}(\Sigma_A, T)\}$ and $\Lambda(\lambda) = P(\langle\lambda, \Phi\rangle)$ for $\lambda \in \mathbb{R}^d$.

Let $D = (D_i)_{i\geq 1}$ be a sequence of digits in the integer basis m. We say that D satisfies property $(\mathcal{P})$ if

Property $(\mathcal{P})$:

(1) The sequence $(\mu_n^D)_{n\geq 1}$ obeys in $\mathbb{R}^d$ the same LDP with rate I as that provided by Theorem 3.3 for $(\mu_n^t)_{n\geq 1}$ (for ν_ψ-almost every t).
(2) The sequence $(\Lambda_n^D)_{n\geq 1}$ satisfies in $\mathbb{R}^d$ the same properties as those provided by Theorem 2.3, Corollary 2.1 and Theorem 2.5 regarding the convergence of $(\Lambda_n^t)_{n\geq 1}$ to Λ (for ν_ψ-almost every t).

Theorem 4.1. *Property $(\mathcal{P})$ implies the normality of $\sum_{i\geq 1} D_i m^{-i}$ in basis m.*

Proof. We prove the equivalent following fact: Let φ be a real valued continuous function on Σ_m. Conjecture 4.1 implies that $\lim_{p\to\infty} S_p\varphi(D)/p = \nu_\psi(\varphi)$.

For $n \geq 1$ let $k_1(n) = (2m)^n$. We have $(n+1)k_1(n+1) \leq 4mnk_1(n)$ for all $n \geq 1$. Fix an integer $n_0 \geq 1$, and for $n \geq n_0$, $1 \leq i \leq 4m-1$ and $0 \leq \ell \leq m^{n_0}$ let $k_{i,\ell}(n) = (i + \ell m^{-n_0})k_1(n)$. The conclusion of Conjecture 4.1(1) holds for every sequence $(k(n))_{n\geq n_0}$ with $k(n) \in \{k_{i,\ell}(n) : 1 \leq i \leq 4m-1,\ 0 \leq \ell \leq m^{n_0}\}$, since $\liminf_{n\to\infty} \log(k(n))/n > \log(m)$.

If p is a positive integer larger than $k_1(1)$, let n_p be the largest integer n such that $nk_1(n) \leq p$. By construction, $p \leq 4mn_pk_1(n_p)$. Let (i_p, ℓ_p) be the unique pair in $\{(i,\ell) : 1 \leq i \leq 4m-1,\ 0 \leq \ell \leq m^{n_0}\}$ such that $n_pk_{i_p,\ell_p}(n_p) \leq p < n_pk_{i_p,\ell_p}(n_p) + n_pk_1(n_p)m^{-n_0}$.

We have

$$\frac{S_p\varphi(D)}{p} = \frac{S_{n_pk_{i_p,\ell_p}(n_p)}}{n_pk_{i_p,\ell_p}(n_p)} + O(m^{-n_0}) \quad (\text{as } p \to \infty),$$

where the constant in $O(m^{-n_0})$ depends only on φ. Consequently, since at fixed n_0 we deal with the finite number of sequences $k(n) \in \{k_{i,\ell}(n) : 1 \leq i \leq 4m-1,\ 0 \leq \ell \leq m^{n_0}\}$, if we prove that Conjecture 4.1 implies that $\lim_{n\to\infty} S_{nk(n)}\varphi(D)/nk(n) = \nu_\psi(\varphi)$ for each such sequence, we will get $\limsup_{p\to\infty} |\frac{S_p\varphi(D)}{p} - \nu_\psi(\varphi)| = O(m^{-n_0})$. Then, letting n_0 tend to ∞ will yield the desired conclusion.

We reduced the problem to showing that $\lim_{n\to\infty} S_{nk(n)}\varphi(D)/nk(n) = \nu_\psi(\varphi)$ whenever $\liminf_{n\to\infty} \log(k(n))/n > \log(m)$. Suppose that $\liminf_{n\to\infty} \log(k(n))/n > \log(m)$. For $\epsilon > 0$, we can write

$$\left|\frac{S_{nk(n)}\varphi(D)}{nk(n)} - \nu_\psi(\varphi)\right| \leq \int_{\mathbb{R}} |x - \nu_\psi(\varphi)|\, \mathrm{d}\mu_n^D(x) \leq \epsilon + 2\|\varphi\|_\infty \mu_n^D(\{x : |x - \nu_\psi(\varphi)| > \epsilon\}),$$

and due to Conjecture 4.1(1), $\mu_n^D(\{x : |x - \nu_\psi(\varphi)| > \epsilon\})$ tends to 0 as $n \to \infty$. Consequently, $\limsup_{n\to\infty} |\frac{S_{nk(n)}\varphi(D)}{nk(n)} - \nu_\psi(\varphi)| \leq \epsilon$ for all $\epsilon > 0$. □

Remark 4.1. One can wonder if, conversely, the normality of $\sum_{i\geq 1} D_i m^{-i}$ implies property $(\mathcal{P})$ for D. To begin with this question, it is interesting to seek an explicit normal number in basis m for which property $(\mathcal{P})$ holds; Champernowne's constant C_m should be investigated.

Our conjecture is the following.

Conjecture 4.1. *For every integer $m \geq 2$, the digits of the fractional part of either Pi or the Euler constant in basis m satisfy $(\mathcal{P})$.*

Conjecture 4.1 is supported by numerical experiments, which focus on the validity of the conclusions of Theorem 2.3 for $(\Lambda_n^D)_{n\geq 1}$. From the numerical point of view, the most tractable situations concern potentials that are constant over the cylinders of the first generation. In the context of digit frequency associated to normality of numbers, it is natural to consider potentials of the form $\Phi_a(t) = \mathbf{1}_{\{a\}}(t_1)$, with $a \in \{0, \dots, m-1\}$. Here, we show simulation results when $m = 10$ and $a = 0$; in this case $\Lambda(\lambda) = \log \frac{9+\exp(\lambda)}{10}$ and $\Lambda^*(x) = x\log(10x) + (1-x)\log(10(1-x)/9)$. We use the 160 millions first decimals of Pi and the Euler constant available at http://www.numberworld.org/constants.html and http://www.ginac.de/~kreckel/news.html.

At first we consider a realization $X_1, \dots, X_N$ of $N = 1.6\cdot 10^8$ independent random variables uniformly distributed in $\{0, \dots, 9\}$, that are viewed as the N first terms of the realization of an infinite sequence of such independent variables $X_1, \dots, X_n, \dots$. In fact these digits are pseudo-random numbers provided by the Mersenne Twister algorithm used in Matlab, so that actually we are also testing how such a sequence really looks like the theoretical one.

At each scale n, we choose a number of intervals $k(n) = \exp(n\Lambda^*(\Lambda'(\lambda_0)))$ with $\lambda_0 = 0.8$, so that $n \cdot k(n) \leq N$ for $n \leq 300$. Due to the fact that $\Phi_0(t)$ depends only on the first digit of t, Λ_n^t is constant over the cylinder $[X_1 \cdots X_{nk(n)}]$ which contains the random sequence $\widetilde{D} = X_1 \cdots X_n \cdots$, and we can estimate it easily.

Let λ_1 and λ_2 the two solutions of the equation $\Lambda^*(\Lambda'(\lambda)) = \Lambda^*(\Lambda'(0.8))$. One has $\lambda_1 \simeq -1.45$ and $\lambda_2 = \lambda_0 = 0.8$.

Fig. 1(a) (left) illustrates the result of Theorem 2.3(1) and (3): the empirical logarithmic moment generating functions $\Lambda_n^{\widetilde{D}}$ converge to the function Λ over the interval (λ_1, λ_2), and on $(-\infty, \lambda_1]$ as well as on $[\lambda_2, \infty)$, $\Lambda - \Lambda_n^{\widetilde{D}}$ converges to Λ translated by an affine map. Fig. 1(a) (right) illustrates the same result in term of the Fenchel–Legendre transform $(\Lambda_n^{\widetilde{D}})^*$, which converges in the interval (x_1, x_2), where $x_1 = \Lambda'(\lambda_1) \simeq 0.0254$ and $x_2 = \Lambda'(\lambda_2) \simeq 0.1983$ (the intervals of convergence are materialized by the dashed blue vertical lines). Moreover, on this figure one observes that the domain over which the functions $(\Lambda_n^{\widetilde{D}})^*$ are finite, which corresponds to $\overline{(\Lambda_n^{\widetilde{D}})'(\mathbb{R})}$, converges to the interval $[x_1, x_2]$. This is predicted by Theorem 2.3(2), since $\overline{(\Lambda_n^{\widetilde{D}})'(\mathbb{R})}$ is equal to the smallest closed interval containing $\{S_n\Phi(T^{n(j-1)}\widetilde{D})/n : 1 \leq j \leq k(n)\}$.

Fig. 1(b) numerically shows that, in terms of the convergence of the logarithmic moment generating functions Λ_n^D and their Fenchel–Legendre transform, the first 160 million decimals of Pi behave exactly like the previous sequence $X_1, \dots, X_N$ (though we do not expose the corresponding figures here, we verified that the same holds for all function Φ_a, $a = 0, \dots, 9$). The same conclusions hold for the 160 millions first decimals of the Euler constant, as shown on Fig. 1(c).

5. Proofs of the main results

Recall that for any $n \geq 1$ and any compact subset B of $\mathcal{D}$, $\delta_n\Lambda(B) = \sup\{|\Lambda(\lambda) - \Lambda_n(\lambda)| : \lambda \in B\}$, $\delta_n\Phi(B) = (\sup_{\lambda\in B}\|\lambda\|)\|S_n\Phi - S_n\Phi_n\|_\infty/n$, $\Lambda^*(B) = \sup\{\Lambda^*(\nabla\Lambda(\lambda)) : \lambda \in B\}$,

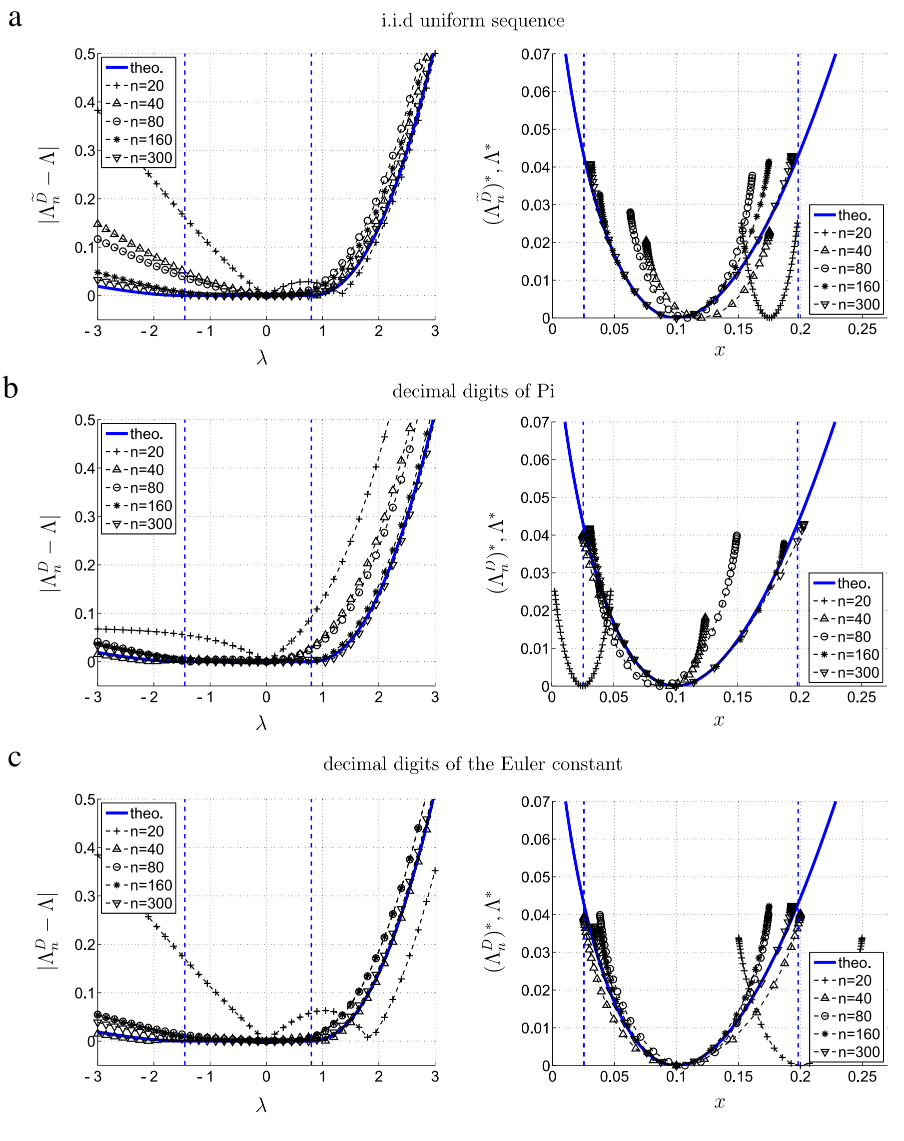

Fig. 1. Behavior of the logarithmic moment generating functions Λ_n^t and their Fenchel–Legendre transform for t equal to an i.i.d sequence of random variables uniformly distributed in $\{0, \ldots, 9\}$ and for t equal to the decimal digits of the number Pi and the Euler constant.

$\xi_1(B) = \sup\{\|\nabla\Lambda(\lambda)\| : \lambda \in B\}$, $\xi_2(B) = \sup\{\frac{1}{2}\,{}^t\lambda \mathrm{D}^2\Lambda(\lambda)\lambda : \lambda \in B\}$, and $\xi(B) = \Lambda^*(B) + \xi_2(B)$.

The following lemma and corollary of its first part will be precious for us. The first part of the lemma can be found in [33], and the second one in [43]. Recall that given a real valued random variable Y defined on $(\Omega, \mathcal{A}, \mathbb{P})$, its quantile function Q_Y is defined as the right-continuous inverse of the tail of $\mathbb{P}_{|Y|}$, the probability distribution of $|Y|$, *i.e.*

$$Q_Y(u) = \inf\{t \geq 0 : \mathbb{P}(|Y| > t) \leq u\} \quad (u \geq 0).$$

Lemma 5.1. (1) *Let* $(Y_j)_{j\geq 1}$ *be a real valued and centered stationary process. For each* $p \in (1,2)$ *and* $N \geq 1$ *one has*

$$\mathbb{E}\left(\left|\sum_{j=1}^{N} Y_j\right|^p\right) \leq C_p\, N \int_0^1 (\alpha_Y^{-1}(u))^{p-1} Q_Y(u)^p \,\mathrm{d}u,$$

with $C_p = 5^p \frac{p(5-2p)}{(p-1)(2-p)}$.

(2) *Let* $(Y_j)_{j\geq 1}$ *be sequence of complex i.i.d. random variables. For each* $p \in (1,2)$ *and* $N \geq 1$ *one has*

$$\mathbb{E}\left(\left|\sum_{j=1}^{N} Y_j\right|^p\right) \leq 2^p\, N \mathbb{E}(|Y_1|^p).$$

Then, the fact that $\int_0^1 Q_Y(u)^p\, du = \mathbb{E}|Y|^p$ together with Hölder's inequality yield

Corollary 5.1. *Let* $(Y_j)_{j\geq 1}$ *be a real valued and centered stationary process. For each* $p \in (1,2)$, $\epsilon > 0$ *and* $N \geq 1$

$$\mathbb{E}\left(\left|\sum_{j=1}^{N} Y_j\right|^p\right) \leq C_p N \left(\int_0^1 (\alpha_Y^{-1}(u))^{(p-1)(1+\epsilon)/\epsilon}\,\mathrm{d}u\right)^{\epsilon/(1+\epsilon)} (\mathbb{E}|Y_1|^{(1+\epsilon)p})^{1/(1+\epsilon)}.$$

We start with the most technical results, namely Theorems 2.3–2.5.

5.1. *Proof of Theorem 2.3*

(1) For $\lambda \in \mathcal{D}$ and $n \geq 1$ we have

$$\Lambda_n^{\omega}(\lambda) = \frac{1}{n} \log \frac{1}{k(n)} \sum_{j=1}^{k(n)} \exp(\langle \lambda, S_n \Phi(T^{n(j-1)} X)\rangle). \tag{5.1}$$

Fix $r_0 > 0$ such that $B(\lambda_0, r_0) \subset \mathcal{D}$. We must prove that we can find $r \in (0, r_0)$ such that almost surely, for all $\lambda \in B(\lambda, r)$, $\Lambda_n^{\omega}(\lambda)$ converges to $\Lambda(\boldsymbol{\lambda})$ as $n \to \infty$.

In order to exploit the mixing properties of the initial process $(X_1, \ldots)$, we use the uniform approximation of $S_n \Phi$ by the functions $S_n \Phi_n$. For $\lambda \in \mathcal{D}$ and $n \geq 1$ let

$$\Lambda_n^{(n),\omega}(\lambda) = \frac{1}{n} \log \frac{1}{k(n)} \sum_{j=1}^{k(n)} \exp(\langle \lambda, S_n \Phi_n(T^{n(j-1)} X)\rangle)$$

and

$$\Lambda_n^{(n)}(\lambda) = \frac{1}{n} \log \mathbb{E} \exp(\langle \lambda, S_n \Phi_n(X)\rangle).$$

By assumption (**A3**) for all $\lambda \in B(\lambda_0, r_0)$ we have

$$\begin{aligned} &\max(|\Lambda_n^{(n),\omega}(\lambda) - \Lambda_n^{\omega}(\lambda)|, |\Lambda_n^{(n)}(\lambda) - \Lambda_n(\lambda)|) \leq \delta_n \Phi(B(\lambda_0, r_0)),\\ &\quad \text{with } \lim_{n\to\infty} \delta_n \Phi(B(\lambda_0, r_0)) = 0. \end{aligned} \tag{5.2}$$

Consequently, it is enough to find $r \in (0, r_0)$ such that almost surely, for all $\lambda \in B(\lambda_0, r)$ we have $\lim_{n\to\infty} \Lambda_n^{(n),\omega}(\lambda) - \Lambda_n^{(n)}(\lambda) = 0$.

For $\lambda \in B(\lambda_0, r_0)$ we write

$$\Lambda_n^{(n),\omega}(\lambda) = \Lambda_n^{(n)}(\lambda) + \frac{1}{n}\log\left(\frac{1}{k(n)}\sum_{j=1}^{k(n)}\exp(\langle\lambda, S_n\Phi_n(T^{n(j-1)}X)\rangle - n\Lambda_n^{(n)}(q))\right),$$

$$= \Lambda_n^{(n)}(\lambda) + \frac{1}{n}\log\left(1 + \frac{1}{k(n)}\sum_{j=1}^{k(n)} Z_j^{(n)}(\lambda)\right), \tag{5.3}$$

where

$$Z_j^{(n)}(\lambda) = \exp(\langle\lambda, S_n\Phi_n(T^{n(j-1)}X)\rangle - n\Lambda_n^{(n)}(\lambda)) - 1. \tag{5.4}$$

Now we notice that $(Z_j^{(n)}(\lambda))_{j\ge 1}$ is stationary and centered, and each $Z_j^{(n)}(\lambda)$ belongs to $\sigma(X_j^{(n)}, X_{j+1}^{(n)})$ (recall (2.2)). Consequently, after writing

$$\mathbb{E}\left(\left|\sum_{j=1}^{k(n)} Z_j^{(n)}(\lambda)\right|^p\right) \le 2^{p-1}\mathbb{E}\left(\left|\sum_{1\le 2j\le k(n)} Z_{2j}^{(n)}(\lambda)\right|^p\right)$$
$$+ 2^{p-1}\mathbb{E}\left(\left|\sum_{1\le 2j+1\le k(n)} Z_{2j+1}^{(n)}(\lambda)\right|^p\right),$$

we can apply Corollary 5.1 and get for $p \in (1, 2)$ and $\epsilon > 0$

$$\mathbb{E}\left(\left|\sum_{j=1}^{k(n)} Z_j^{(n)}(\lambda)\right|^p\right) \le M(n, p, \epsilon)k(n)(\mathbb{E}|Z_1^{(n)}(\lambda)|^{(1+\epsilon)p})^{1/(1+\epsilon)}, \tag{5.5}$$

where

$$M(n, p, \epsilon) = 2^{p-1}C_p\left(\int_0^1 (\alpha_{X^{(n)}}^{-1}(u))^{(p-1)(1+\epsilon)/\epsilon}\,\mathrm{d}u\right)^{\epsilon/(1+\epsilon)}.$$

From now on we fix $p \in (1, 2)$ close enough to one and $\epsilon_0 > 0$ small enough so that for all $\lambda \in B_0 = B(\lambda_0, r_0/2)$ and $\epsilon \in (0, \epsilon_0)$ we have $p(1+\epsilon)\lambda \in \widetilde{B}_0 = B(\lambda_0, r_0)$.

We have (using successively the convexity of $u \ge 0 \mapsto u^{(1+\epsilon)p}$ and the subadditivity of $v \ge 0 \mapsto v^{1/(1+\epsilon)}$ to get the second and third lines)

$$(\mathbb{E}|Z_1^{(n)}(\lambda)|^{(1+\epsilon)p})^{1/(1+\epsilon)} \le 2^{p(1+\epsilon)-1}(\mathbb{E}[\exp((1+\epsilon)p\langle\lambda, S_n\Phi_n(T^{n(j-1)}X)\rangle$$
$$- n(1+\epsilon)p\Lambda_n^{(n)}(\lambda))] + 1)^{1/(1+\epsilon)}$$
$$\le 2^p\left(1 + \exp\left[n\left(\frac{\Lambda_n^{(n)}(p(1+\epsilon)\lambda)}{1+\epsilon} - p\Lambda_n^{(n)}(\lambda)\right)\right]\right)$$
$$\le 2^p\left(1 + \exp\left[n\left(\frac{\Lambda(p(1+\epsilon)\lambda)}{1+\epsilon} - p\Lambda(\lambda) + 3\delta_n(\Lambda, \Phi)(\widetilde{B}_0)\right)\right]\right), \tag{5.6}$$

where we have used (**A1**) and (**A3**). Since Λ is differentiable at λ_0, by using the first order Taylor expansion of Λ at λ_0, for each $r \in (0, r_0/2)$, uniformly in $\lambda \in B(\lambda_0, r)$ we have

$$\frac{\Lambda(p(1+\epsilon)\lambda)}{1+\epsilon} - p\Lambda(\lambda) = (1-p)(\Lambda(\lambda_0) - \langle\lambda_0, \nabla\Lambda(\lambda_0)\rangle) + \xi(\epsilon)\epsilon$$

$$\begin{aligned}&+\eta(p,r)(p-1+r)\\&=(p-1)\Lambda^*(\nabla\Lambda(\lambda_0))+\xi(\epsilon)\epsilon+\eta(p,r)(p-1+r),\end{aligned}$$

where $\xi(\epsilon)$ is bounded over $(0,\epsilon_0)$ and $\eta(p,r)$ tends to 0 as p tends to 1^+ and r tends to 0^+. This yields

$$\begin{aligned}(\mathbb{E}|Z_1^{(n)}(\lambda)|^{(1+\epsilon)p})^{1/(1+\epsilon)} &\le 2^p(1+\exp[n((p-1)\Lambda^*(\nabla\Lambda(\lambda_0))\\&\quad+\xi(\epsilon)\epsilon+\eta(p,r)(p-1+r)+3\delta_n(\Lambda,\Phi)(\widetilde{B}_0))])\\&\le 2^{p+1}\exp[n((p-1)\Lambda^*(\nabla\Lambda(\lambda_0))\\&\quad+\xi(\epsilon)\epsilon+\eta(p,r)(p-1+r)+3\delta_n(\Lambda,\Phi)(\widetilde{B}_0))],\end{aligned}$$

the last inequality coming from the fact that $\Lambda^*\ge 0$.

Recall (**A2**). Let $h=(p-1)(1+\epsilon)/\epsilon$ and notice that $\alpha_{X^{(n)}}^{-1}\le\alpha_X^{-1}$. This yields

$$M(n,p,\epsilon)\le 2^{p-1}C_pM_h^{\epsilon/(1+\epsilon)}<\infty. \tag{5.7}$$

Thus, due to (5.5), for n large enough so that $3\delta_n(\Lambda,\Phi)(\widetilde{B}_0)\le\epsilon$, uniformly in $\lambda\in B(\lambda_0,r)$ we have

$$\begin{aligned}\mathbb{E}\left(\left|\sum_{j=1}^{k(n)}Z_j^{(n)}(\lambda)\right|^p\right)&\le 2^{2p}C_pM_h^{\epsilon/(1+\epsilon)}k(n)\exp[n((p-1)\Lambda^*(\nabla\Lambda(\lambda))\\&\quad+\widetilde{\xi}(\epsilon)\epsilon+\eta(p,r)(p-1+r))],\end{aligned}$$

where $\widetilde{\xi}(\epsilon)=\xi(\epsilon)+1$, hence

$$\begin{aligned}\mathbb{P}\left(\left|\frac{1}{k(n)}\sum_{j=1}^{k(n)}Z_j^{(n)}(\lambda)\right|>\epsilon\right)&\le\epsilon^{-p}k(n)^{-p}\mathbb{E}\left(\left|\sum_{j=1}^{k(n)}Z_j^{(n)}(\lambda)\right|^p\right)\\&\le 2^{2p}C_pM_h^{\epsilon/(1+\epsilon)}\epsilon^{-p}k(n)^{1-p}\exp[n((p-1)\Lambda^*(\nabla\Lambda(\lambda))\\&\quad+\widetilde{\xi}(\epsilon)\epsilon+\eta(p,r)(p-1+r))].\end{aligned} \tag{5.8}$$

Let $\eta>0$ such that $k(n)\ge\exp(n(\Lambda^*(\nabla\Lambda(\lambda_0))+\eta))$ for n large enough. The previous inequality yields, for n large enough, uniformly in $\lambda\in B(\lambda_0,r)$,

$$\begin{aligned}&\mathbb{P}\left(\left|\frac{1}{k(n)}\sum_{j=1}^{k(n)}Z_j^{(n)}(\lambda)\right|>\epsilon\right)\\&\quad\le 2^{2p}C_pM_h^{\epsilon/(1+\epsilon)}\epsilon^{-p}\exp(n((1-p)\eta+\widetilde{\xi}(\epsilon)\epsilon+\eta(p,r)(p-1+r))).\end{aligned}$$

Hence, fixing p close enough to 1 and r small enough so that $\eta(p,r)(p-1+r)\le(p-1)\eta/2$, we get

$$\mathbb{P}\left(\left|\frac{1}{k(n)}\sum_{j=1}^{k(n)}Z_j^{(n)}(\lambda)\right|>\epsilon\right)\le 2^{2p}C_pM_h^{\epsilon/(1+\epsilon)}\epsilon^{-p}\exp(n((1-p)\eta/2+\widetilde{\xi}(\epsilon)\epsilon)).$$

Then, for every ϵ small enough so that $\widetilde{\xi}(\epsilon)\epsilon\le(p-1)\eta/4$ we have

$$\sum_{n\ge1}\mathbb{P}\left(\left|\frac{1}{k(n)}\sum_{j=1}^{k(n)}Z_j^{(n)}(\lambda)\right|>\epsilon\right)<\infty$$

for every $\lambda \in B(\lambda_0, r)$ (notice that at fixed p, h tends to ∞ as ϵ tends to 0, this is why we need **(A2)**). Now, by the Borel–Cantelli lemma and (5.3), we can conclude that for every $\lambda \in B(\lambda_0, r)$ we have $\lim_{n\to\infty} \Lambda_n^{(n),\omega}(\lambda) - \Lambda_n^{(n)}(\lambda) = 0$ almost surely, hence $\lim_{n\to\infty} \Lambda_n^{\omega}(\lambda) = \Lambda(\lambda)$ almost surely. From this we deduce that, with probability 1, $\lim_{n\to\infty} \Lambda_n^{\omega}(\lambda) = \Lambda(\lambda)$ for every λ in a countable and dense subset of $B(\lambda_0, r)$. Since the functions $\Lambda_n^{\omega}(\lambda)$ and Λ are convex, we deduce from Theorem 10.8 in [34] that almost surely, $\Lambda(\lambda)$ converges to $\Lambda(\lambda)$ for all λ in $B(\lambda_0, r)$.

(2) The first part is a direct consequence of (1) and Theorem 1.5.

Now let $x_0 = \nabla\Lambda(\lambda_0)$. Let $\eta > 0$ such that $k(n) \le \exp(n(\Lambda^*(x_0) - \eta))$ for n large enough. For ϵ small enough, if n is large enough, we have

$$\begin{aligned}\mathbb{P}\left(\exists\, 1 \le j \le k(n) : \frac{S_n \Phi(T^{(j-1)n} X)}{n} \in B(x_0, \epsilon)\right) &\le k(n)\mathbb{P}\left(\frac{S_n \Phi(X)}{n} \in B(x_0, \epsilon)\right) \\ &\le \exp(-n\eta/2)\end{aligned}$$

by Theorem 1.5. Thus, by the Borel–Cantelli lemma we see that if ϵ is small enough, with probability 1, for n large enough $\{1 \le j \le k(n) : \frac{S_n \Phi(T^{(j-1)n} X)}{n} \in B(x_0, \epsilon)\}$ is empty.

(3) Let $P : t \ge 0 \mapsto \Lambda(t\lambda_0)$. Since P is strictly convex at 1, there exists $\epsilon \in (0, 1)$ such that P' exists and is not constant over $[1 - \epsilon, 1]$, so that $P^*(P'(t)) < P^*(P'(1)) = \Lambda^*(x_0)$ for $t \in [1 - \epsilon, 1]$. Consequently, if we set $P_n^{\omega}(t) = \Lambda_n^{\omega}(t\lambda_0)$, we deduce from Theorem 2.3(1) that with probability 1, for all $\eta \in (0, \epsilon)$, $P_n^{\omega}(1 - \eta)$ converges to $P(1 - \eta)$ as n tends to ∞, and the same holds for their derivatives (by convexity). Now, we notice that for any $s > 1$, by the superadditivity of $y \ge 0 \mapsto y^s$, we have $P_n^{\omega}(s(1 - \eta)) \le sP_n^{\omega}(1 - \eta) + (s - 1)\log(k_n)/n$. Thus, due to our assumption, $\limsup_{n\to\infty} P_n^{\omega}(s(1 - \eta)) \le sP(1 - \eta) + (s - 1)\Lambda^*(x_0)$. If $t > 1$, for each $\eta \in (0, \epsilon)$, if we set $s = t/(1 - \eta)$, we get $\limsup_{n\to\infty} P_n^{\omega}(t) \le tP(1 - \eta)/(1 - \eta) + (t - 1 + \eta)\Lambda^*(x_0)/(1 - \eta)$. Consequently, $\limsup_{n\to\infty} P_n^{\omega}(t) \le tP(1) + (t - 1)\Lambda^*(x_0) = \Lambda(\lambda_0) + (t - 1)\langle\lambda_0, x_0\rangle$.

On the other hand, by convexity, for all $n \ge 1$ and $\eta \in (0, 1)$, for $t > 1$ we have $P_n^{\omega}(t) \ge P_n^{\omega}(1-\eta) + (t-1+\eta)(P_n^{\omega})'(1-\eta)$. Thus $\liminf_{n\to\infty} P_n^{\omega}(t) \ge P(1-\eta) + (t-1+\eta)(P)'(1-\eta)$, and letting η go to 1, we get $\liminf_{n\to\infty} P_n^{\omega}(t) \ge P(1) + (t-1)(P)'(1) = \Lambda(\lambda_0) + (t-1)\langle\lambda_0, x_0\rangle$. Thus we have the conclusion.

5.2. *Proof of Theorem 2.4*

Since Λ is twice continuously differentiable, by using the second order Taylor expansion of Λ we can get for all $\lambda \in B_{\rho/2}$ and for all $p \in (1, 2)$ and $\epsilon > 0$ such that $p(1 + \epsilon)B_{\rho/2} \subset B_\rho$

$$\frac{\Lambda(p(1+\epsilon)\lambda)}{1+\epsilon} - p\Lambda(\lambda) = \frac{(p(1+\epsilon) - 1)\Lambda^*(\nabla\Lambda(\lambda)) + \delta(p, \epsilon)}{1+\epsilon},$$

where $|\delta(p, \epsilon)| \le \xi_2(B_\rho)(p(1 + \epsilon) - 1)^2$. Consequently, for all $\lambda \in B_{\rho/2}$ and for all $p \in (1, 2)$ and $\epsilon \in (0, 1/2)$ such that $p(1 + \epsilon)B_{\rho/2} \subset B_\rho$,

$$\begin{aligned}&\left|\frac{\Lambda(p(1+\epsilon)\lambda)}{1+\epsilon} - p\Lambda(\lambda) - (p-1)\Lambda^*(\nabla\Lambda(\lambda))\right| \le \frac{\Lambda^*(\nabla\Lambda(\lambda))\epsilon + \delta(p, \epsilon)}{1+\epsilon} \\ &\quad\le \Lambda^*(\nabla\Lambda(\lambda))\epsilon + \xi_2(B_\rho)(p-1)^2 + (p^2\epsilon + 2p(p-1))\xi_2(B_\rho)\epsilon \\ &\quad\le \xi_2(B_\rho)(p-1)^2 + (p^2\epsilon + 2p(p-1) + 1)(\xi_1(B_\rho) + \xi_2(B_\rho))\epsilon \\ &\quad= \xi_2(B_\rho)(p-1)^2 + (p^2\epsilon + 2p(p-1) + 1)\xi(B_\rho)\epsilon.\end{aligned}$$

Thus, for p close enough to 1 and ϵ close enough to 0,

$$\left|\frac{\Lambda(p(1+\epsilon)\lambda)}{1+\epsilon} - p\Lambda(\lambda) - (p-1)\Lambda^*(\nabla\Lambda(\lambda))\right| \leq \xi_2(B_\rho)(p-1)^2 + 2\xi(B_\rho)\epsilon.$$

Let $(k(n))_{n\geq 1}$ and $(\epsilon_n)_{n\geq 1}$ be as in the statement, and take $\epsilon = \epsilon_n$ and $p = p_n = 1 + \sqrt{\epsilon_n}$. Defining the variables $Z_j^{(n)}$ as in the proof of Theorem 2.3, by using (5.6) and (5.8) we can get

$$\mathbb{P}\left(\left|\frac{1}{k(n)}\sum_{j=1}^{k(n)} Z_j^{(n)}(\lambda)\right| > \epsilon_n\right) \leq 2^{2p_n} C_{p_n} M_{h_n}^{\epsilon_n/(1+\epsilon_n)} \epsilon_n^{-p_n} \exp(-n\tau(n,\lambda)) \tag{5.9}$$

with $h_n = (p_n - 1)(1+\epsilon_n)/\epsilon_n$ and $\tau(n,\lambda) = (p_n - 1)(\log(k(n))/n - \Lambda^*(\nabla\Lambda(\lambda)) - \xi_2(B_\rho)(p_n - 1)^2 - 2\xi(B_\rho)\epsilon_n - 3\delta_n\Lambda(B_\rho) - 3\delta_n\Phi(B_\rho))$. We have (recall the value of C_p given in Lemma 5.1(1))

$$2^{2p_n} C_{p_n} \epsilon_n^{-p_n} = O(\epsilon_n^{-3/2-\sqrt{\epsilon_n}}) = O(\epsilon_n^{-3/2})$$

as n tends to ∞. Moreover,

$$\tau(n,\lambda) \geq \tau(n) = \sqrt{\epsilon_n}(\log(k(n))/n - \Lambda^*(B)) - 3(\xi(B_\rho)\epsilon_n + \delta_n\Lambda(B_\rho) + \delta_n\Phi(B_\rho)),$$

and an estimation provided at the end of this proof shows that $M_{h_n}^{\epsilon_n/(1+\epsilon_n)} = O(1)$ as n tends to ∞. Thus,

$$\mathbb{P}\left(\left|\frac{1}{k(n)}\sum_{j=1}^{k(n)} Z_j^{(n)}(\lambda)\right| > \epsilon_n\right) = O(\epsilon_n^{-3/2}\exp(-n\tau(n))).$$

Now, let $g(n) = \lfloor\log_2(\sqrt{d}/\epsilon_n)\rfloor + 1$ and $\mathcal{G}_n(B_{\rho/2}) = \{(k_1,\ldots,k_d) \in \mathbb{Z}^d : (k_1 2^{-g(n)},\ldots, k_d 2^{-g(n)}) \in B_{\rho/2}\}$. There exists a constant $C(B_{\rho/2})$ depending on the volume of $B_{\rho/2}$ only such that $\#\mathcal{G}_n(B_{\rho/2}) \leq C(B_{\rho/2})\epsilon_n^{-d}$, hence

$$\mathbb{P}\left(\exists\lambda \in \mathcal{G}_n(B_{\rho/2}) : \left|\frac{1}{k(n)}\sum_{j=1}^{k(n)} Z_j^{(n)}(\lambda)\right| > \epsilon_n\right) = O(\epsilon_n^{-(3/2+d)}\exp(-n\tau(n))),$$

and due to (2.7), the Borel–Cantelli lemma ensures that, with probability 1, for n large enough, for all $\lambda \in \mathcal{G}_n(B_{\rho/2})$, $\left|\frac{1}{k(n)}\sum_{j=1}^{k(n)} Z_j^{(n)}(\lambda)\right| \leq \epsilon_n$. This can be used in (5.3) and combined with (5.2) to get for n large enough

$$\sup_{\lambda_n \in \mathcal{G}_n(B_{\rho/2})} |\Lambda_n^\omega(\lambda_n) - \Lambda(\lambda_n)| \leq (\epsilon_n + \epsilon_n^2)/n + \delta_n\Lambda(B_\rho) + \delta_n\Phi(B_\rho). \tag{5.10}$$

Since $(\mathcal{G}_n(B_{\rho/2}))_{n\geq 1}$ is increasing and $\bigcup_{n\geq 1}\mathcal{G}_n(B_{\rho/2})$ is dense in $B_{\rho/2}$, the convexity of Λ_n^ω and Λ ensures that Λ_n^ω converges uniformly to Λ over $B_{\rho/2}$, and $\nabla\Lambda_n^\omega$ converges uniformly to $\nabla\Lambda$ over B (see [34], Th. 10.8 and 25.7). Thus, for any $\eta > 0$, if n is large enough, we have both (5.10) and $\sup_{\lambda\in B}\|\nabla\Lambda_n^\omega - \nabla\Lambda\| \leq \eta/2$, so that for all $\lambda \in B$, we can choose $\lambda_n \in \mathcal{G}_n(B)$ such that $\|\lambda - \lambda_n\| \leq \epsilon_n$, hence

$$\begin{aligned}
|\Lambda_n^\omega(\lambda) - \Lambda(\lambda)| &\leq |\Lambda_n^\omega(\lambda_n) - \Lambda(\lambda_n)| + |\Lambda_n^\omega(\lambda) - \Lambda_n^\omega(\lambda_n)| + |\Lambda(\lambda) - \Lambda(\lambda_n)| \\
&\leq (\epsilon_n + \epsilon_n^2)/n + \delta_n\Lambda(B_\rho) + \delta_n\Phi(B_\rho) + (\eta/2 + 2\max_{\lambda'\in B}\|\nabla\Lambda\|)\|\lambda - \lambda_n\| \\
&\leq (\eta + 2\max_{\lambda'\in B}\|\nabla\Lambda\|)\epsilon_n + \delta_n\Lambda(B_\rho) + \delta_n\Phi(B_\rho).
\end{aligned}$$

It remains to prove that $M_{h_n}^{\epsilon_n/(1+\epsilon_n)} = O(1)$ as n tends to ∞. Due to (**A2′**), there exists $C > 0$ such that $\alpha_X(u)^{-1} \le C|\log(u)|^{1/\theta}$ for all $u \in (0, 1]$. This yields for $h > 0$

$$\begin{aligned} M_h &= \int_0^1 (\alpha_X(u)^{-1})^h \,\mathrm{d}u \le \int_0^1 C^h |\log(u)|^{h/\theta} \,\mathrm{d}u \\ &= C^h \Gamma(1 + h/\theta) = O(C^h (N(h,\theta)/e)^{N(h,\theta)} \sqrt{2\pi N(h,\theta)}), \end{aligned}$$

where $N(h,\theta) = \lfloor h/\theta \rfloor + 1$ and we have use Stirling's formula.

Now, we can use the fact that $h_n = (1 + \epsilon_n)/\sqrt{\epsilon_n}$ and the estimate above to conclude that $M_{h_n}^{\epsilon_n/(1+\epsilon_n)} = O(1)$ as n tends to ∞.

5.3. *Proof of Theorem 2.5*

Here we have $\delta_n \Lambda = \delta_n \Phi = 0$, so that with respect to the proof of Theorem 2.4, we can consider the centered, independent and identically distributed variables

$$Z_{n,j}(\lambda) = \exp(\langle \lambda, S_n \Phi(T^{nj} X)\rangle - n\Lambda(\lambda)) - 1,$$

instead of the $Z_j^{(n)}(\lambda)$, with $\Lambda(\lambda) = \log \mathbb{E}(\exp\langle \lambda, X\rangle)$. Now we can use Lemma 5.1(2) instead of Lemma 5.1(1). This yields, for p small enough so that $pB \subset B_\rho$ and $\lambda \in B$

$$\begin{aligned} \mathbb{E}\left(\left|\sum_{j=1}^{k(n)} Z_{n,j}(\lambda)\right|^p\right) &\le 2^p k(n)(\mathbb{E}|Z_{n,1}(\lambda)|^p) \\ &\le 2^{2p-1} k(n)(1 + \mathbb{E}\exp(p\langle \lambda, S_n \Phi(T^{nj} X)\rangle - np\Lambda(\lambda))) \\ &= 2^{2p-1} k(n)(1 + \exp(n(\Lambda(p\lambda) - p\Lambda(\lambda)))) \\ &\le 2^{2p} k(n) \exp(n[(p-1)\Lambda^*(\nabla\Lambda(\lambda)) + (p-1)^2 \xi_2(B_\rho)]). \end{aligned}$$

The proof finishes as that of Theorem 2.4(1).

5.4. *Proof of Theorem 2.1*

(1) Fix $\theta > 0$. For $n \ge 1$, we denote by $\mu_n^{(n)}$ the probability distribution of $S_n \Phi_n(X)/n$ and by $\mu_n^{(n),\omega}$ the empirical distribution of $(S_n \Phi_n(T^{(j-1)} X(\omega)))/n$.

Recall that $(\delta_n)_{n\ge1}$ is defined in (**A3**). Let $B_n \in \{B(x, r-\delta_n), B(x, r+\delta_n)\}$. We can estimate $\mathbb{P}\big(|\mu_n^{(n),\omega}(B_n) - \mu_n^{(n)}(B_n)| \ge \theta\mu_n^{(n)}(B_n)\big)$ as when we get (5.5) thanks to Corollary 5.1.

Fix $p \in (1, 2)$, $\epsilon > 0$ and $h = (p-1)(1+\epsilon)/\epsilon$. We have

$$\begin{aligned} &\mathbb{P}(|\mu_n^{(n),\omega}(B_n) - \mu_n^{(n)}(B_n)| \ge \theta\mu_n^{(n)}(B_n)) \\ &\le (\theta\mu_n^{(n)}(B_n))^{-p} \mathbb{E}\left(\frac{1}{k(n)^p}\left|\sum_{j=1}^{k(n)} \mathbf{1}_{B_n}(S_n \Phi_n(T^{(j-1)n} X)/n) - \mathbb{P}(S_n \Phi_n(X)/n \in B_n)\right|^p\right) \\ &\le 2^{p-1} C_p M_h^{\epsilon/(1+\epsilon)} (\theta\mu_n^{(n)}(B_n))^{-p} k(n)^{1-p} \mathbb{E}(|\mathbf{1}_{B_n}(S_n \Phi_n(X)) \\ &\quad - \mathbb{P}(S_n \Phi_n(X) \in B_n)|^{p(1+\epsilon)})^{1/(1+\epsilon)} \\ &\le 2^{p-1} C_p M_h^{\epsilon/(1+\epsilon)} (\theta\mu_n^{(n)}(B_n))^{-p} k(n)^{1-p} \\ &\quad \times (2^{p(1+\epsilon)-1} (\mathbb{P}(S_n \Phi_n(X) \in B_n) + \mathbb{P}(S_n \Phi_n(X) \in B_n)^{p(1+\epsilon)})^{1/(1+\epsilon)}) \end{aligned}$$

$$\leq 2^{2p-1} C_p M_h^{\epsilon/(1+\epsilon)} \theta^{-p} k(n)^{1-p} \mu_n^{(n)}(B_n)^{-p} (\mu_n^{(n)}(B_n)^{1/(1+\epsilon)} + \mu_n^{(n)}(B_n)^p)$$
$$\leq 2^{2p} C_p M_h^{\epsilon/(1+\epsilon)} \theta^{-p} k(n)^{1-p} \mu_n^{(n)}(B_n)^{-p+1/(1+\epsilon)}$$
$$\leq 2^{2p} C_p M_h^{\epsilon/(1+\epsilon)} \theta^{-p} k(n)^{1-p} \mu_n(B(x, r-2\delta_n))^{-p+1/(1+\epsilon)}.$$

Now suppose that $\liminf_{n\to\infty} \frac{\log k(n)}{n} > I(x)$ and let $\eta > 0$ such that $k(n) \geq \exp(n(I(x)+\eta))$ for n large enough. Since $(\mu_n)_{n\geq 1}$ satisfies the LDP with rate function I, for n large enough we have $\mu_n(B(x, r-2\delta_n)) \geq \exp(-n(I(x)+\eta/4))$. Consequently, we can choose ϵ small enough so that $k(n)^{1-p} \mu_n(B(x, r-2\delta_n))^{-p+1/(1+\epsilon)} \leq \exp(-n(p-1)\eta/2)$ for n large enough. Then, the previous bound for $\mathbb{P}\big(|\mu_n^{(n),\omega}(B_n) - \mu_n^{(n)}(B_n)| \geq \theta \mu_n^{(n)}(B_n)\big)$ yields $\sum_{n\geq 1} \mathbb{P}\big(|\mu_n^{(n),\omega}(B_n) - \mu_n^{(n)}(B_n)| \geq \theta \mu_n^{(n)}(B_n)\big) < \infty$ for any $\theta > 0$. Hence, by the Borel–Cantelli lemma we get that with probability one, $\lim_{n\to\infty} \frac{1}{n} \log \frac{\mu_n^{(n),\omega}(B_n)}{\mu_n^{(n)}(B_n)} = 0$ for $B_n \in \{B(x, r-\delta_n), B(x, r+\delta_n)\}$.

Moreover, we have $\mu_n^{(n),\omega}(B(x, r-\delta_n)) \leq \mu_n^{\omega}(B(x,r)) \leq \mu_n^{(n),\omega}(B(x, r+\delta_n))$, and on the other hand we have $\mu_n(B(x, r-2\delta_n)) \leq \mu_n^{(n)}(B(x, r-\delta_n)) \leq \mu_n(B(x,r)) \leq \mu_n^{(n)}(B(x, r+\delta_n)) \leq \mu_n(B(x, r+2\delta_n))$.

Consequently, for any $r > 0$, with probability 1,

$$\begin{aligned}\liminf_{n\to\infty} -\frac{1}{n} \log \mu_n(B(x, r+2\delta_n)) &\leq \liminf_{n\to\infty} -\frac{1}{n} \log \mu_n^{\omega}(B(x,r)) \\ &\leq \limsup_{n\to\infty} -\frac{1}{n} \log \mu_n^{\omega}(B(x,r)) \\ &\leq \limsup_{n\to\infty} -\frac{1}{n} \log \mu_n(B(x, r-2\delta_n)).\end{aligned}$$

This implies that with probability 1, for all $r \in \mathbb{Q}_+^*$ we have

$$\begin{aligned}&\liminf_{n\to\infty} -\frac{1}{n} \log \mu_n(B(x, 3r/2)) \leq \liminf_{n\to\infty} -\frac{1}{n} \log \mu_n^{\omega}(B(x,r)) \\ &\quad \leq \limsup_{n\to\infty} -\frac{1}{n} \log \mu_n^{\omega}(B(x,r)) \leq \limsup_{n\to\infty} -\frac{1}{n} \log \mu_n(B(x, r/2)).\end{aligned}$$

But since $(\mu_n)_{n\geq 1}$ satisfies the LDP with rate function I, we have for all $y \in \mathcal{Y}$ (see [11], Th. 4.1.18)

$$\lim_{s\to 0^+} -\liminf_{n\to\infty} \frac{1}{n} \log \mu_n(B(y,s)) = \lim_{s\to 0^+} -\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(B(y,s)) = I(y). \tag{5.11}$$

This, together with the previous inequalities yields the desired result.

At last, suppose that $\limsup_{n\to\infty} \frac{\log k(n)}{n} < I(x)$. An estimate similar to that used to establish the second part of Theorem 2.3(2) yields the desired result.

(2) Let $B \in B_{\mathcal{Y}}$ and $\gamma > 0$. If $\mu_n(B) > 0$ we have

$$\begin{aligned}&\mathbb{P}(\mu_n^{\omega}(B) > \exp(n\gamma)\mu_n(B)^{1-1/n}) \\ &\quad \leq \mathbb{P}\left(\sum_{j=1}^{k(n)} \mathbf{1}_B(S_n \Phi(T^{(j-1)n} X)) > k(n) \exp(n\gamma) \mu_n(B)^{1-1/n}\right)\end{aligned}$$

$$\le k(n)^{-1}\exp(-n\gamma)\mu_n(B)^{1/n-1}\mathbb{E}\left(\sum_{j=1}^{k(n)}\mathbf{1}_B(S_n\Phi(T^{(j-1)n}X))\right)$$
$$=\exp(-n\gamma)\mu_n(B)^{1/n},$$

and clearly if $\mu_n(B)=0$ then $\mu_n^\omega(B)=0$ almost surely so that we also have $\mathbb{P}(\mu_n^\omega(B)>\exp(n\gamma)\mu_n(B)^{1-1/n})\le\exp(-n\gamma)\mu_n(B)^{1/n}$.

Now fix $r>0$ and take $B=B(x,r)$. Since $\sum_{n\ge1}\exp(-n\gamma)\mu_n(B)^{1/n}<\infty$, the Borel–Cantelli lemma yields $\limsup_{n\to\infty}\frac{1}{n}\log\mu_n^\omega(B(x,r))\le\gamma+\limsup_{n\to\infty}\frac{1}{n}\log\mu_n(B(x,r))$ almost surely. This holds for all $\gamma>0$, so $\limsup_{n\to\infty}\frac{1}{n}\log\mu_n^\omega(B(x,r))\le\limsup_{n\to\infty}\frac{1}{n}\log\mu_n(B(x,r))$ almost surely. This enough to conclude thanks to (5.11) and the fact that $I(x)=\infty$.

(3) Let $\alpha<\infty$ and $K_\alpha\subset\mathcal{Y}$, a compact set such that $\limsup_{n\to\infty}\frac{1}{n}\log\mu_n(K_\alpha^c)\le-2\alpha$. By using the estimate obtained above with $B=K_\alpha^c$ and $\gamma=\alpha$ we get that with probability 1, $\limsup_{n\to\infty}\frac{1}{n}\log\mu_n^\omega(K_\alpha^c)\le-\alpha$.

5.5. *Proof of Theorem 2.2*

Our goal is to prove that, with probability 1, for all $y\in\mathcal{Y}$ we have

$$\lim_{r\to0^+}\liminf_{n\to\infty}-\frac{1}{n}\log\mu_n^\omega(B(y,r))=\lim_{r\to0^+}\limsup_{n\to\infty}-\frac{1}{n}\log\mu_n^\omega(B(y,r))=I(y). \tag{5.12}$$

Then, due to Theorem 4.1.11 in [11], we have the desired almost sure weak LDP.

Let $\mathcal{D}$ be a dense countable subset of $\mathcal{D}_I$. We can deduce from the end of the proof of Theorem 2.1(1) that there exists a measurable subset Ω' of Ω such that $\mathbb{P}(\Omega')=1$ and for all $\omega\in\Omega'$, for all $x\in\mathcal{D}$ and for all $r\in\mathbb{Q}_+^*$ we have

$$\liminf_{n\to\infty}-\frac{1}{n}\log\mu_n(B(x,3r/2))\le\liminf_{n\to\infty}-\frac{1}{n}\log\mu_n^\omega(B(x,r))$$
$$\le\limsup_{n\to\infty}-\frac{1}{n}\log\mu_n^\omega(B(x,r))\le\limsup_{n\to\infty}-\frac{1}{n}\log\mu_n(B(x,r/2)).$$

Now let $y\in\mathcal{D}_I$. For all $s>0$ we can find $x\in\mathcal{D}$ as well as a rational number $0<r<s$ such that $B(y,s/4)\subset B(x,r/2)\subset B(y,s)\subset B(x,3/2r)\subset B(y,2s)$. Consequently, for all $\omega\in\Omega'$, $y\in\mathcal{D}_I$ and $r>0$ we have

$$\liminf_{n\to\infty}-\frac{1}{n}\log\mu_n(B(y,2s))\le\liminf_{n\to\infty}-\frac{1}{n}\log\mu_n^\omega(B(y,s))$$
$$\le\limsup_{n\to\infty}-\frac{1}{n}\log\mu_n^\omega(B(y,s))\le\limsup_{n\to\infty}-\frac{1}{n}\log\mu_n(B(y,s/4)).$$

Due to (5.11), for all $\omega\in\Omega'$ and $y\in\mathcal{D}_I$ we get

$$\lim_{s\to0^+}\liminf_{n\to\infty}-\frac{1}{n}\log\mu_n^\omega(B(y,s))=\lim_{s\to0^+}\limsup_{n\to\infty}-\frac{1}{n}\log\mu_n^\omega(B(y,s))=I(y),$$

that is (5.12) for $y\in\mathcal{D}_I$.

Now suppose that $\mathcal{Y}\setminus\mathcal{D}_I\neq\emptyset$ and let $\mathcal{D}'$ be a dense subset of $\mathcal{Y}\setminus\mathcal{D}_I$. Due to the facts established in the proof of Theorem 2.1(2), there exists a measurable subset Ω' of Ω

such that $\mathbb{P}(\Omega') = 1$ and for all $\omega \in \Omega'$, for all $x \in \mathcal{D}'$, for all $r \in \mathbb{Q}_+^*$ we have $\limsup_{n\to\infty} \frac{1}{n} \log \mu_n^\omega(B(x,r)) \le \limsup_{n\to\infty} \frac{1}{n} \log \mu_n(B(x,r))$.

Now, for all $y \in \mathcal{Y} \setminus \mathcal{D}_I$ and $s > 0$, we can find $x \in \mathcal{D}'$ and $0 < s < r \in \mathbb{Q}$ such that $B(y,s) \subset B(x,r) \subset B(y,2s)$, and the previous inequality yields, for all $\omega' \in \Omega'$, $\limsup_{n\to\infty} \frac{1}{n} \log \mu_n^\omega(B(y,s)) \le \limsup_{n\to\infty} \frac{1}{n} \log \mu_n(B(y,2s))$. This yields (5.12).

6. Proofs of Theorems 3.1 and 3.2

6.1. Proof of Theorem 3.1

Each interval $J_{k(n),j}$ can be decomposed into a union of n consecutive closed intervals $J_{k(n),j,i}$ of length $1/nk(n)$. The increments $\Delta W(J_{k(n),j,i})$ take the form $(nk(n))^{-1/2} X_{k(n),j,i}$, where $(X_{k(n),j,i})_{\substack{1\le j\le k(n)\\ 1\le i\le \kappa(n)}}$ is a family of $nk(n)$ centered Gaussian vectors of covariance matrix the identity. Thus $\Delta W(J_{k(n),j}) = (nk(n))^{-1/2} S_n(j)$ with $S_n(j) = \sum_{i=1}^n X_{k(n),j,i}$. Let

$$Z_{n,j}(\lambda) = \exp(\langle \lambda, S_n(j)\rangle - n\Lambda(\lambda)) - 1,$$

with $\Lambda(\lambda) = \log \mathbb{E}(\exp\langle \lambda, X_{k(n),j,i}\rangle) = \|\lambda\|^2/2$, hence $\Lambda^*(\nabla\Lambda(\lambda)) = \|\lambda\|^2/2$ and ${}^t\lambda \mathrm{D}^2\Lambda(\lambda)\lambda = \|\lambda\|^2$ for all $\lambda \in \mathbb{R}^d$. As in the proof of Theorem 2.5 we have

$$\mathbb{E}\left(\left|\sum_{j=1}^{k(n)} Z_{n,j}(\lambda)\right|^p\right) \le 2^{2p-1} k(n)(1 + \exp(n(\Lambda(p\lambda) - p\Lambda(\lambda))))$$

which, due to the special form of Λ, yields

$$\mathbb{E}\left(\left|\sum_{j=1}^{k(n)} Z_{n,j}(\lambda)\right|^p\right) \le 2^{2p} k(n) \exp(n[(p-1)\|\lambda\|^2/2 + (p-1)^2\|\lambda\|^2/2]).$$

Then, we can use the same approach as that used in the proof of Theorem 2.4 to get that under (3.1), with probability 1,

$$\lim_{n\to\infty} \left(\Lambda_n^\omega(\lambda) = \frac{1}{n} \log \frac{1}{k(n)} \sum_{j=1}^{k(n)} \exp(\langle \lambda, S_n(j)\rangle)\right) = \Lambda(\lambda) = \|\lambda\|^2/2$$

for a dense and countable subset of points $\lambda \in B$, hence for all $\lambda \in B$ by convexity of the functions Λ_n^ω. This is enough to get the result.

6.2. Proof of Theorem 3.2

We leave the reader adapt the lines of the proof of Theorem 2.2 to the present situation. The only change is that here for each $n \ge 1$ one must consider the i.i.d sequence of Brownian motions obtained by juxtaposition of the $k(n)$ sequences of n Brownian motions $((W_{k(n),j,i})_{t\in[0,1]})_{1\le i\le n}$, $1 \le j \le k(n)$, where $W_{k(n),j,i}(t) = (nk(n))^{1/2}(W(\frac{j-1}{k(n)} + \frac{i-1+t}{nk(n)}) - W(\frac{j-1}{k(n)} + \frac{i-1}{nk(n)}))$, so that $W_{k(n),j}/n^{1/2} = S_n(j)/n$ with $S_n(j) = \sum_{i=1}^n W_{k(n),j,i}$.

Acknowledgment

The authors were partially supported by the French National Research Agency Project "DMASC".

References

[1] D.H. Bailey, P.B. Borwein, S. Plouffe, On the rapid computation of various polylogarithmic constants, Math. Comput. 66 (1997) 903–913.
[2] D.H. Bailey, R.E. Crandall, On the random character of fundamental constant expansions, Experiment. Math. 10 (2001) 175–190.
[3] V. Baladi, Positive Transfer Operators and Decay of Correlations, in: Advanced Series in Nonlinear Dynamics, vol. 16, World Scientific, 2000.
[4] J. Barral, A.H. Fan, Covering numbers of different points in Dvoretzky covering, Bull. Sci. Math. 119 (2005) 275–317.
[5] J. Barral, D.J. Feng, Weighted thermodynamic formalism and applications. arXiv:math/0909.4247v1.
[6] J. Barral, B.B. Mandelbrot, Multifractal products of cylindrical pulses, Probab. Theory Relat. Fields 124 (2002) 409–430.
[7] P. Billingsley, Convergence of Probability Measures, 2nd edition, John Wiley & Sons, New York, 1999.
[8] P. Bougerol, J. Lacroix, Products of Matrices with Applications to Schrödinger Operators, Birkhäuser, 1985.
[9] R. Bowen, Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms, in: Lecture Notes in Mathematics, vol. 470, Springer, 1975.
[10] L. Carleson, On the support of harmonic measure for sets of Cantor type, Ann. Acad. Sci. Fenn. Math. 10 (1985) 113–123.
[11] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, Jones and Bartlett Publishers, 1993.
[12] A. Dvoretzky, On covering a circle by randomly placed arcs, Proc. Natl. Acad. Sci. USA 42 (1956) 199–203.
[13] R.S. Ellis, Entropy, Large Deviations, and Statistical Mechanics, Springer-Verlag, 1985.
[14] A.-H. Fan, How many intervals cover a point in Dvoretzky covering?, Israël J. Math. 131 (2002) 157–184.
[15] A.H. Fan, D.-J. Feng, On the distribution of long-term time averages on symbolic space, J. Statist. Phys. 99 (3–4) (2000) 813–856.
[16] A.H. Fan, D.-J. Feng, J. Wu, Recurrence, dimension and entropy, J. London Math. Soc. (2) 64 (1) (2001) 229–244.
[17] A.-H Fan, L.-M. Liao, B.-W. Wang, J. Wu, On Khintchine exponents and Lyapounov exponents of continued fractions, Ergodic Theory Dynam. Systems 29 (2009) 73–109.
[18] D.-J. Feng, The variational principle for products of non-negative matrices, Nonlinearity 17 (2004) 447–457.
[19] D.-J. Feng, Lyapunov exponents for products of matrices and multifractal analysis. Part II: general matrices, Israël J. Math. 170 (2009) 355–394.
[20] D.-J. Feng, W. Huang, Lyapunov spectrum of asymptotically sub-additive potentials, Comm. Math. Phys., doi:10.1007/s00220-010-1031-x.
[21] B. Hasselblatt, A. Katok, Introduction to the Modern Theory of Dynamical Systems, Cambridge University Press, 1995.
[22] J.-P. Kahane, Some Random Series of Functions, Cambridge University Press, 1985.
[23] J.-P. Kahane, J. Peyrière, Sur certaines martingales de Benoît Mandelbrot, Adv. Math. 22 (1976) 131–145.
[24] Q. Liu, A. Rouault, On two measures defined on the boundary of a branching tree, in: Classical and Modern Branching Processes (Minneapolis, MN, 1994), in: IMA Vol. Math. Appl., vol. 84, Springer-Verlag, New York, 1997, pp. 187–201.
[25] R. Lyons, Random walks, capacity and percolation on trees, Ann. Probab. 20 (1992) 2043–2088.
[26] R. Lyons, R. Pemantle, Y. Peres, Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes, Ann. Probab. 23 (1995) 1125–1138.
[27] N.G. Makarov, Fine structure of harmonic measure, St. Peterburg Math. J. 10 (2) (1999) 217–268.
[28] B.B. Mandelbrot, Intermittent turbulence in self-similar cascades: divergence of hight moments and dimension of the carrier, J. Fluid Mech. 62 (1974) 331–358.
[29] B.B. Mandelbrot, On Dvoretzky coverings for the circle, Z. Wahrsch. Verw. Geb. 22 (1972) 158–160.
[30] A.A. Mogulskii, Large deviations for processes with independent increments, Ann. Probab. 21 (1993) 202–215.
[31] S. Orey, S. Pelikan, Large deviation principle for principles for stationary processes, Ann. Probab. 16 (1988) 1481–1495.
[32] S. Orey, S. Pelikan, Deviations of trajectory averages and the defect in Pesin's formula for Anosov diffeomorphisms, Trans. Amer. Math. Soc. 315 (1989) 741–753.
[33] E. Rio, Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants, Springer, 2000.
[34] R.T. Rockafellar, Convex Analysis, in: Princeton Mathematical Series, vol. 28, Princeton University Press, Princeton, NJ, 1970.

[35] D. Ruelle, Thermodynamic Formalism, in: Encyclopedia of Mathematics and its Applications, vol. 5, Addison-Wesley, 1978.
[36] D. Ruelle, Thermodynamic formalism for maps satisfying positive expansiveness and specification, Nonlinearity 5 (1992) 1223–1236.
[37] O. Sarig, Subexponential decay of correlations, Invent. Math. 150 (2002) 629–653.
[38] L.A. Shepp, Covering the line with random intervals, Z. Wahrsch. Verw. Geb. 23 (1972) 163–170.
[39] L.A. Shepp, Covering the circle with random arc, Israel J. Math. 11 (1972) 328–345.
[40] D.W. Strook, An Introduction to the Theory of Large Deviations, Springer-Verlag, 1984.
[41] F. Takens, E. Verbitskiy, On the variational principle for the topological entropy of certain non-compact sets, Ergodic Theory Dynam. Systems 23 (2003) 317–348.
[42] S.R.S. Varadhan, Large Deviations and Applications, in: SIAM CBMS-NSF Regional Conf. Series, SIAM, 1984.
[43] E. von Bahr, C.G. Esseen, Inequalities for the rth absolute moment of a sum of random variables, $1 \leqq r \leqq 2$, Ann. Math. Statist. 36 (1965) 299–303.
[44] L.-S. Young, Some large deviations results for dynamical systems, Trans. Amer. Math. Soc. 318 (1990) 525–541.