WEAK GIBBS MEASURES: SPEED OF CONVERGENCE TO ENTROPY, TOPOLOGICAL AND GEOMETRICAL ASPECTS

PAULO VARANDAS AND YUN ZHAO

Abstract. In this paper we obtain exponential large deviation bounds in the Shannon-McMillan-Breiman convergence formula for entropy in the case of weak Gibbs measures and topologically mixing subshifts of finite type. We also prove almost sure estimates for the error term in the convergence to entropy given by Shannon-McMillan-Breiman formula for both uniformly and non-uniformly expanding shifts. Finally we establish a topological characterization of large deviations bounds for Gibbs measures and deduce some of their topological and geometrical aspects: the local entropy is zero and the topological pressure of positive measure sets is total. Some applications include large deviation estimates for Lyapunov exponents, pointwise dimension and slow return times.

1. Introduction

Since it was introduced in dynamical systems more than fifty years ago, entropy has become an important ingredient in the characterization of the complexity of dynamical systems in both topological and measure theoretical senses.

From the topological viewpoint, the topological entropy reflects the topological complexity of the dynamical system and is a topological invariant. In other words, since any two conjugated systems have the same topological entropy, distinct topological entropy becomes a criterium to detect elements in different C^0-conjugacy classes. From the ergodic viewpoint, the metric entropy of invariant measures turns out to be a surprisingly universal concept in ergodic theory since it appears in the study of different subjects as information theory, Poincaré recurrence, and in the analysis of either local or global complexities. Just as an illustration of its universal nature, metric entropy is characterized as the exponential growth rate of: the measure of decreasing partition elements and/or dynamical balls, the number of dynamical balls and partition elements necessary to cover a relevant part of the phase space, or the recurrence rate to elements of a given partition (see Section 3 for the precise formulas).

Topological and metric entropy are two fundamental concepts in ergodic theory and are related by the variational principle for entropy (see e.g. [11]). Maximal entropy measures, in case they exist, are a special class of measures whose metric complexity reflect the topological complexity of the dynamics. In many cases, in which maximal entropy measures arise from the thermodynamical formalism,

[0]2000 *Mathematics Subject classification*: 37M25 37A30, 37D35, 37H15, 60F10,

Date: January 21, 2016.

Key words and phrases. Weak Gibbs measures, large deviations, asymptotically additive and sub-additive potentials, non-additive thermodynamical formalism, law of iterated logarithm, capacity pressure.

maximal entropy measures are proved to satisfy some Gibbs property: the measure reflects physically the dynamical system under the interaction of a potential (see Section 2 for the precise definition). For very complete accounts on entropy we refer the reader to [6, 11] and references therein. Starting at the end of the eighties with the work of Falconer, a detailed study of the thermodynamic formalism associated to non-additive sequences of functions was developed. The idea is to generalize classical results in the field replacing the pressure of a continuous function with that of a sequence of functions. This theory was originally developed to study dimension theory of non-conformal repellers but in recent years it has found applications in different settings.

Since entropy plays such an interesting role in the characterization of complexity in dynamics and information theory, it is an important question to know how fast this value can be computed, that is, to estimate the error term and convergence rate in its several limit expressions. The primary goal of this paper is to provide almost sure and convergence in probability estimates for the error term in the Shannon-McMillan-Breiman formula for the entropy of weak Gibbs measures. This class of measures is physically relevant and it is widely known that these appear naturally in the context of equilibrium states in the thermodynamical formalism of non-uniformly hyperbolic transformations. We prove a large deviations principle for the convergence to entropy of weak Gibbs measures, thus exponentially fast convergence in probability, whenever the measure of cylinders decreases with distinct asymptotic exponential rates in different regions of the phase space (see Theorems A and B for precise statements). We also estimate the error term of the almost sure convergence to entropy in the case of non-uniformly expanding shifts (see Theorem C). To the best of our knowledge, these convergence estimates are new even in the context of Gibbs measures for uniformly expanding maps. The second goal of this paper is to provide a topological and geometrical characterization of deviation sets. In rough terms, under some extra conditions, the topological pressure of the deviation sets is strictly smaller than the topological pressure if and only if exponential large deviations hold. This was motivated by [13, 2, 3] and the precise statement is given in Theorem D. From the geometrical viewpoint, we deduce that any positive measure set (w.r.t. the weak Gibbs measure) carries capacity pressure equal to the topological pressure (Corollary A) and that the local entropy is almost everywhere constant to zero (Corollary B).

This paper is organized as follows. In Section 2 we recall definitions necessary. Section 3 is devoted to the statements and proofs of the results concerning convergence to entropy while Section 4 is devoted to the proofs of the results relating large deviations and topological pressure of deviation sets and their geometrical and topological consequences. Finally in Section 5 we give some examples and applications.

2. Preliminaries

In this section we recall the necessary definitions in the theory of thermodynamical formalism for continuous dynamical systems.

2.1. **Non-additive sequences of potentials.**

Let $f : X \to X$ be a continuous transformation on a compact metric space X with the metric d, and let $C(X)$ be the Banach space of all continuous functions from X to $\mathbb{R}$ equipped with the supremum norm $\|\cdot\|$. Denote by $\mathcal{M}_f$ and $\mathcal{E}_f$ the set of all f-invariant respectively, ergodic f-invariant Borel probability measures on X. Given a sequence of continuous potentials $\Phi = \{\varphi_n\} \subset C(X)^{\mathbb{N}}$ we say that Φ is:

- *sub-additive*, if $\varphi_{m+n} \leq \varphi_m + \varphi_n \circ f^m$ for every $m, n \geq 1$;
- *almost additive*, if there exists a uniform constant $C > 0$ such that $\varphi_m + \varphi_n \circ f^m - C \leq \varphi_{m+n} \leq \varphi_m + \varphi_n \circ f^m + C$ for every $m, n \geq 1$; and
- *asymptotically additive*, if for any $\xi > 0$ there exists a continuous function φ_ξ such that $\limsup_{n\to\infty} \frac{1}{n} \|\varphi_n - S_n\varphi_\xi\| < \xi$, where $S_n\varphi_\xi := \sum_{j=0}^{n-1} \varphi_\xi \circ f^j$ denotes the usual Birkhoff sum of the function φ_ξ.

For short, we shall say that a sequence of continuous potentials $\Phi = \{\varphi_n\} \subset C(X)^{\mathbb{N}}$ is *non-additive* if it satisfies any of the previous three conditions.

2.2. **Topological pressure.** A subset $E \subset X$ is called (n, ϵ)-*separated* if all distinct $x, y \in E$ satisfy $d_n(x, y) := \max\{d(f^i x, f^i y) : 0 \leq i < n\} > \epsilon$. Given a non-additive sequence of continuous potentials $\Phi = \{\varphi_n\}$ on X, the *topological pressure of f with respect to* Φ is defined by

$$P_{\text{top}}(f, \Phi) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \sup_{E_n} \{Z_n(\Phi, E_n, \epsilon)\} \tag{2.1}$$

where $Z_n(\Phi, E_n, \epsilon) = \sum_{y \in E_n} \exp(\varphi_n(y))$ and the supremum is taken over all (n, ϵ)-separated sets. The following variational principle relates the non-additive topological pressure with the natural modifications of the measure-theoretic free energy. If $\Phi = \{\varphi_n\}$ is a non-additive potential (namely sub-additive, almost additive or asymptotically additive) on X then

$$P_{\text{top}}(f, \Phi) = \sup\{h_\mu(f) + \mathcal{F}_*(\mu, \Phi) : \mu \in \mathcal{M}_f,\ \mathcal{F}_*(\mu, \Phi) \neq -\infty\}.$$

Here $\mathcal{F}_*(\mu, \Phi) := \lim\limits_{n\to\infty} \frac{1}{n} \int \varphi_n \, d\mu$ and $h_\mu(f)$ is the metric entropy of f with respect to the measure μ (see [21] for the detailed definition). An f-invariant probability measure μ that attains the supremum is called an *equilibrium state* for f with respect to Φ. We refer the reader to [1, 5, 7, 15] for the proof of this variational principle and details on topological pressure of non-additive potentials.

2.3. **Weak Gibbs measures.** In many situations the equilibrium states arise as invariant measures absolutely continuous with respect to weak Gibbs measures, that we now recall. For notational simplicity, we use $P(\Phi)$ instead of $P_{\text{top}}(f, \Phi)$ to denote the topological pressure of f with respect to Φ when no confusion is possible.

Definition 2.1. *Given a sequence of continuous functions $\Phi = \{\varphi_n\} \subset C(X)^{\mathbb{N}}$, we say that a probability measure ν is a* weak Gibbs measure with respect to Φ on $\Lambda \subset X$, *if the set Λ has full ν-measure and for every $x \in \Lambda$ there exists a sequence of positive constants $\{K_n(x)\}_{n\geq 1}$ satisfying $\lim\limits_{n\to\infty} \frac{1}{n} \log K_n(x) = 0$ and for every $n \geq 1$*

$$K_n^{-1}(x) \leq \frac{\nu(B_n(x, \epsilon))}{e^{-nP(\Phi) + \varphi_n(x)}} \leq K_n(x)$$

where $B_n(x, \epsilon) := \{y \in X : d_n(x, y) < \epsilon\}$ denotes the dynamical ball centered at x of radius ϵ and of length n. We say that ν is a Gibbs measure with respect to Φ *if there exists $K > 0$ such that the previous property holds with $K_n = K$ (independent of n and x).*

Given a Hölder continuous potential $\varphi : X \to \mathbb{R}$, we say that μ_φ is a weak Gibbs measure for f with respect to φ if the condition of Definition 2.1 holds for sequence $\Phi = \{\varphi_n\}$ where $\varphi_n = S_n\varphi$, in this case we denote the topological pressure with respect to Φ by $P(\varphi)$. The uniform version of Gibbs measure for almost additive potentials above corresponds to the one used by Barreira [1] and Mummert [15] in the uniformly expanding setting. In the case of additive potentials, weak Gibbs measures appear naturally in non-uniformly hyperbolic dynamics (see e.g. [18]). In most of the applications through the article, we will be interested in establishing ergodic properties for invariant measures with the weak Gibbs property.

2.4. **Bounded distortion property.** An ingredient to prove that some Gibbs properties holds is to obtain a bounded distortion property for the Jacobian of the measure.

Definition 2.2. *We say that a sequence of continuous functions $\Phi = \{\varphi_n\}$ satisfies the* bounded distortion property *(resp.* weak bounded distortion property*), if*

$$\sup_{n\in\mathbb{N}} \gamma_n(\Phi,\delta) < \infty \quad \textit{for some } \delta > 0 \quad \Big(\textit{respectively } \limsup_{\delta\to 0}\limsup_{n\to\infty} \frac{\gamma_n(\Phi,\delta)}{n} = 0\Big),$$

where $\gamma_n(\Phi,\delta) = \sup_{x\in X}\{|\varphi_n(y) - \varphi_n(z)| : \ y, z \in B_n(x,\delta)\}$.

The later conditions are sometimes referred as strong and weak *Bowen conditions*, respectively. These hold, e.g. for almost-additive potentials $\Psi = \{\psi_n\}$ (see [22]). We observe that given a partition $\mathcal{P}$ of X, if $\mathcal{P}^{(n)}(x)$ denotes the element of the refined partition $\mathcal{P}^{(n)} := \bigvee_{i=0}^{n-1} f^{-i}\mathcal{P}$ which contains x, analogous notions can be defined by replacing the dynamical ball $B_n(x,\delta)$ by $\mathcal{P}^{(n)}(x)$. In the case of hyperbolic dynamics, it is known that sequences of almost-additive potentials $\Psi = \{\psi_n\}$ admit a weak Gibbs measure. If, in addition, the almost additive potential satisfies the bounded distortion property, then it has a unique equilibrium state which coincides with the Gibbs measure with respect to Φ (see [1, 15] for the proof). Thus, invariant probability measures that are weak Gibbs do exist in this context.

3. Velocity of convergence to entropy

The main results of this section are exponential large deviations estimates for the convergence to entropy given by Shannon-McMillan-Breiman theorem (Theorem A and Theorem B) and, an almost sure estimate for the error term in the convergence (Theorem C) in the context of one-sided topologically mixing subshifts of finite type and weak Gibbs measures. In Section 5 we illustrate how these shifts can be used to model other classes of non-uniformly expanding maps. To the best of our knowledge, these results are new even in the stronger context of Gibbs measures in uniformly expanding dynamics.

3.1. **Convergence to entropy.** The notion of metric entropy is one fundamental tool to describe the measure theoretical complexity of a dynamical system. In particular, this justifies that it is a very well studied concept in ergodic theory and its connection with other relevant dynamical quantities as dynamical balls, recurrence estimates or Lyapunov exponents is of great interest.

3.1.1. *Entropy formulas.* Given a measurable map $f : X \to X$ on a compact metric space (X, d) and an f-invariant ergodic probability measure μ, recall that the *metric entropy* $h_\mu(f)$ of f with respect to μ is defined by

$$h_\mu(f) = \sup_\xi h_\mu(f, \xi)$$

where ξ is a finite partition of X and $h_\mu(f, \xi)$ is the metric entropy of f with respect to the partition ξ. Moreover, if one considers the first return time to n-cylinders given by $R_n(x, \xi) = \inf\{k \geq 1 : f^k(x) \in \xi^{(n)}(x)\}$ then $h_\mu(f, \xi)$ can be computed by

$$h_\mu(f, \xi) = \lim_{n\to\infty} -\frac{1}{n} \log \mu(\xi^{(n)}(x)) \quad \text{for } \mu\text{- a.e. } x \qquad \text{(Shannon-McMillan-Breiman)}$$

and also

$$h_\mu(f, \xi) = \lim_{n\to\infty} \frac{1}{n} \log R_n(x, \xi) \quad \text{for } \mu\text{- a.e. } x. \qquad \text{(Ornstein-Weiss)}$$

The previous expressions are very useful for computing the metric entropy when ξ is a generating partition, that is, $h_\mu(f) = h_\mu(f, \xi)$. Recall that by the Kolmogorov-Sinai criterium, every partition ξ is a generating partition for μ provided that $\bigvee_{i=0}^{+\infty} f^{-i}\xi$ is μ-almost everywhere the partition into singletons (see e.g. [6, 21]).

Alternatively, the entropy $h_\mu(f)$ can be defined in terms of dynamical balls. Taking $R_n(x, \epsilon) = \inf\{k \geq 1 : f^k(x) \in B_n(x, \epsilon)\}$ one can compute the metric entropy by

$$h_\mu(f) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} -\frac{1}{n} \log \mu(B_n(x, \epsilon)) \quad \text{for } \mu\text{- a.e. } x \qquad \text{(Brin-Katok)}$$

and

$$h_\mu(f) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log R_n(x, \epsilon) \quad \text{for } \mu\text{- a.e. } x \qquad \text{(Ornstein-Weiss like)}$$

We refer the reader to [11, 19, 6, 17] and references therein for other entropy formulas.

Throughout this section we assume (X, f) to be a topologically mixing subshift of finite type, where $X = \Sigma_A^+ \subset \{1, \ldots, p\}^{\mathbb{N}}$ is associated to a transition matrix $A \in \mathcal{M}_{p\times p}(\mathbb{R})$, and let $f = \sigma : \Sigma_A^+ \to \Sigma_A^+$ be the shift map. Let Σ_A^+ be endowed with the distance

$$d(\omega, \omega') = 2^{-k}, \tag{3.1}$$

where $\omega = (\omega_1\omega_2 \ldots)$, $\omega' = (\omega_1'\omega_2' \ldots)$ and $k = \min\{n \geq 1 : \omega_n \neq \omega_n'\}$. Let $C_n(x) = \{y \in \Sigma_A^+ : y_i = x_i,\ 0 \leq i < n\}$ denote a cylinder of length n that contains the point x and let $\mathcal{P} = \{[0], [1], \cdots, [p-1]\}$ be the natural partition of Σ_A^+ by cylinders of size one.

3.1.2. *Non-additive sequences arising in entropy.* It is natural to address questions on the velocity at which the previous convergences hold. The velocity can be expressed both in terms of large deviations and almost sure estimates. The later convergences may be very distinct depending on the type of invariant measures (e.g. Dirac measures for which convergence is immediately attained). Moreover some difficulties arise from the fact that the expressions in the previous limit formulas are not monotone nor additive in general. Our starting point is the following result.

Proposition 3.1. *Let $\Phi = \{\varphi_n\}$ be an almost-additive sequence of continuous potentials on Σ_A^+ and assume that μ_Φ is a fully supported Gibbs measure for σ with respect to Φ. Then the sequence of potentials $\Psi = \{-\log \mu_\Phi(\mathcal{P}^{(n)}(x))\}$ are continuous and almost-additive.*

Proof. Since μ_Φ has full support, the functions in Ψ are well defined and finite everywhere. Since the partition functions $\Sigma_A^+ \to \mathcal{P}$ given by $x \mapsto \mathcal{P}^{(n)}(x)$ are locally constant, every potential in the family $\Psi = \{-\log \mu_\Phi(\mathcal{P}^{(n)}(x))\}$ is continuous. By the Gibbs property for μ_Φ, there exists $K > 0$ so that

$$\frac{1}{K} \leq \frac{\mu_\Phi(\mathcal{P}^{(n)}(x))}{e^{-P(\Phi)n+\varphi_n(x)}} \leq K$$

for every $n \geq 1$ and every $x \in \Sigma_A^+$, where $P(\Phi) = P_{\text{top}}(\sigma, \Phi)$. In consequence, using that the sequence $\Phi = \{\varphi_n\}$ is almost-additive there exists $C > 0$ so that for every $n \geq 1$ and $x \in \Sigma_A^+$

$$\begin{aligned}
\psi_{m+n}(x) &= -\log \mu_\Phi(\mathcal{P}^{(m+n)}(x)) \\
&\leq \log K + (m+n)P(\Phi) - \varphi_{m+n}(x) \\
&\leq \log K + mP(\Phi) + nP(\Phi) - \varphi_n(x) - \varphi_m(\sigma^n(x)) + C \\
&\leq \psi_n(x) + \psi_m(\sigma^n(x)) + 3\log K + C.
\end{aligned}$$

Since a similar lower estimate is completely analogous, we deduce that Ψ is an almost-additive sequence of continuous functions. This finishes the proof of the proposition. $\square$

Remark 3.1. *Let us mention that additivity property for the previous sequence Ψ can also follow from the independence for μ. If μ is Bernoulli then*

$$\begin{aligned}
\log \mu(\mathcal{P}^{(m+n)}(x)) &= \log \mu(\mathcal{P}^{(n)}(x) \cap \sigma^{-n}(\mathcal{P}^{(m)}(\sigma^n(x)))) \\
&= \log \mu(\mathcal{P}^{(n)}(x)) + \log \mu(\sigma^{-n}(\mathcal{P}^{(m)}(\sigma^n(x)))) \\
&= \log \mu(\mathcal{P}^{(n)}(x)) + \log \mu(\mathcal{P}^{(m)}(\sigma^n(x)))
\end{aligned}$$

and so $\psi_{m+n} = \psi_n + \psi_m \circ \sigma^n$ (e.g. maximal entropy measure μ_0 for the full shift Σ_p^+ in which case $\mu_0(\mathcal{P}^{(n)}(x)) = p^{-n}$ for every $n \geq 1$, c.f. [4]).

One does not expect the stronger almost-additivity property in the case that μ_Φ is a weak Gibbs measure even in the case when the constants $K_n(\cdot)$ do not depend on x as we now explain. In fact, assume $\Phi = \{\varphi_n\}$ is an almost-additive sequence of potentials on Σ_A^+ and μ_Φ is a full supported probability measure so that there are constants $K_n \geq 1$ for which $\limsup_{n\to\infty} \frac{1}{n} \log K_n = 0$ and

$$\frac{1}{K_n} \leq \frac{\mu_\Phi(\mathcal{P}^{(n)}(x))}{e^{-P(\Phi)n+\varphi_n(x)}} \leq K_n \tag{3.2}$$

for every $n \geq 1$ and every $x \in \Sigma_A^+$. The sequence $\Psi = \{-\log \mu_\Phi(\mathcal{P}^{(n)}(x))\}$, provided $\Phi = \{\varphi_n\}$ is almost additive, verifies the following: there exists $C > 0$ so that for every $n \geq 1$ and $x \in \Sigma_A^+$

$$\begin{aligned}
\psi_{m+n}(x) &= -\log \mu_\Phi(\mathcal{P}^{(m+n)}(x)) \\
&\leq \log K_{m+n} + (m+n)P(\Phi) - \varphi_{m+n}(x) \\
&\leq \log K_{m+n} + P(\Phi)m + P(\Phi)n - \varphi_n(x) - \varphi_m(\sigma^n(x)) + C \\
&\leq \psi_n(x) + \psi_m(\sigma^n(x)) + \log K_{m+n} + \log K_m + \log K_n + C.
\end{aligned}$$

Without loss of generality, we may assume that $(K_n)_n$ is a non-decreasing sequence and so, since a similar lower estimate is completely analogous, we deduce that

$$|\psi_{m+n}(x) - \psi_n(x) - \psi_m(\sigma^n(x))| \leq 4 \log K_{m+n} \tag{3.3}$$

for every $m, n \geq 1$. Motivated by equation(3.3) it is natural to ask if a family of continuous functions $\Psi = \{\psi_n\}$ admits constants $C_n \geq 0$ satisfying $\limsup_{n\to\infty} \frac{1}{n} C_n = 0$ and $\psi_n(x) + \psi_m(\sigma^n(x)) - C_{m+n} \leq \psi_{m+n}(x) \leq \psi_n(x) + \psi_m(\sigma^n(x)) + C_{m+n}$ for every $m, n \geq 1$ and $x \in \Sigma_A^+$, is necessarily asymptotically additive. We can provide a partial answer to the above question provided the constants C_n satisfy some growth condition as follows.

Proposition 3.2. *Let $\Psi = \{\psi_n\}$ on Σ_A^+ be a sequence of continuous potentials and $(C_n)_{n\geq 1}$ be a sequence of non-negative constants for which*

$$\psi_n(x) + \psi_m(\sigma^n(x)) - C_{m+n} \leq \psi_{m+n}(x) \leq \psi_n(x) + \psi_m(\sigma^n(x)) + C_{m+n} \tag{3.4}$$

for every $m, n \geq 1$ and $x \in \Sigma_A^+$. If we assume that

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} C_j < +\infty \tag{3.5}$$

then $\Psi = \{\psi_n\}$ is an asymptotically additive sequence of potentials on Σ_A^+.

Proof. Fix a positive integer k. For each natural number n, we write $n = ks + r$, where $s \geq 0$, $0 \leq r < k$. Then, for any integer $0 \leq j < k$ we have

$$\psi_n(x) \leq \psi_j(x) + \sum_{l=0}^{s-2} \psi_k(\sigma^{kl}\sigma^j x) + \psi_{k+r-j}(\sigma^{k(s-1)}\sigma^j x) + C_n + \sum_{l=0}^{s-2} C_{n-lk-j}.$$

where $\psi_0(x) \equiv 0$. Let $M_1 = \max_{j=1,\cdots 2l} \|\psi_j\|$. Adding $\psi_n(x)$ when j takes all the natural values from 0 to $k-1$, we have

$$k\psi_n(x) \leq 2kM_1 + \sum_{i=0}^{(s-1)k-1} \psi_k(\sigma^i x) + kC_n + \sum_{i=k+r+1}^{n-1} C_i.$$

Hence

$$\begin{aligned}
\psi_n(x) &\leq 2M_1 + \sum_{i=0}^{(s-1)k-1} \frac{1}{k}\psi_k(\sigma^i x) + C_n + \frac{1}{k} \sum_{i=k+r+1}^{n-1} C_i \\
&\leq 4M_1 + \sum_{i=0}^{n-1} \frac{1}{k}\psi_k(\sigma^i x) + C_n + \frac{1}{k} \sum_{i=k+r+1}^{n-1} C_i
\end{aligned}$$

Similarly, we can prove that

$$\psi_n(x) \geq -4M_1 + \sum_{i=0}^{n-1} \frac{1}{k}\psi_k(\sigma^i x) - C_n - \frac{1}{k}\sum_{i=k+r+1}^{n-1} C_i.$$

Since $A := \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} C_j < +\infty$ and, consequently, $\limsup_{n\to\infty} \frac{1}{n} C_n = 0$, this yields that

$$\limsup_{n\to\infty} \frac{1}{n}\|\psi_n - S_n(\frac{1}{k}\psi_k)\| \leq \frac{A}{k},$$

which implies that the sequence $\Psi = \{\psi_n\}$ is asymptotically additive. $\square$

Some comments concerning the proposition are in order. If the sequence $(C_n)_n$ is bounded then the sequence $\Psi = \{\psi_n\}$ is almost additive, hence asymptotically additive. In the case that sequence $(C_n)_n$ is non-decreasing, equation (3.5) implies that there exists $K > 0$ so that

$$\frac{1}{2}C_n = \frac{1}{2n} nC_n \leq \frac{1}{2n}\sum_{j=n}^{2n-1} C_j \leq \frac{1}{2n}\sum_{j=0}^{2n-1} C_j \leq K$$

for all large n, from which we conclude once more that $(C_n)_n$ is bounded. So, the previous proposition yields new results in the case of unbounded and non-monotone sequences $(C_n)_n$ satisfying (3.5).

The following simple lemma answers a converse of the previous question.

Lemma 3.1. *Let $\Phi = \{\varphi_n\}$ be an asymptotically additive potential on Σ_A^+. Then, for any $\varepsilon > 0$ there exists $N_\varepsilon \geq 1$ so that $\|\varphi_{n+m} - \varphi_n - \varphi_m \circ \sigma^n\| \leq 2(n+m)\varepsilon$ for every $m, n \geq N_\varepsilon$.*

Proof. Assume that $\Phi = \{\varphi_n\}$ is an asymptotically additive potential. Then, for any $\varepsilon > 0$ there exists $\varphi_\varepsilon \in C(\Sigma_A^+)$ and $N_\varepsilon \geq 1$ so that $\|\varphi_n - S_n\varphi_\varepsilon\| < \varepsilon n$ for every $n \geq N_\varepsilon$. In particular, if $n, m \geq N_\varepsilon$ then

$$\begin{aligned}
\|\varphi_{n+m} - \varphi_n - \varphi_m \circ \sigma^n\| &\leq \|\varphi_{n+m} - S_{n+m}\varphi_\varepsilon\| + \|\varphi_n - S_n\varphi_\varepsilon\| \\
&\quad + \|\varphi_m \circ \sigma^n - S_m\varphi_\varepsilon \circ \sigma^n\| \\
&\leq \varepsilon(n+m) + \varepsilon n + \varepsilon m = 2(n+m)\varepsilon,
\end{aligned}$$

proving the lemma. $\square$

Remark 3.2. *Let us observe that more recently Iommi and Yayama* [9, Theorem 2.1] *established that if μ_Φ is a weak Gibbs measure in the sense of* (3.2) *for an almost additive sequence Φ, then the sequence $\Psi = \{-\log \mu_\Phi(\mathcal{P}^{(n)}(x))\}$ is necessarily asymptotically additive.*

3.1.3. *Large deviations principles.* We now proceed to provide large deviations estimates for the convergence to entropy in the case of Shannon-McMillan-Breiman's entropy formula. We make use of the large deviations results obtained in [20] for sub-additive and asymptotically additive continuous potentials.

Given any σ-invariant probability measure η and an almost additive sequence of continuous functions $\Phi = \{\varphi_n\}$, assume that μ_Φ is a full supported Gibbs measure for σ with respect to Φ. Consider the almost additive sequence of continuous observables $\Psi = \{\psi_n\}$ given by $\psi_n = -\log \mu_\Phi(\mathcal{P}^{(n)}(x))$ (c.f. Proposition 3.1). By the L^1-convergence in Kingman's sub-additive ergodic theorem ([12]) we have

$$\mathcal{F}_*(\eta, \Psi) := \lim_{n\to\infty} \int -\frac{1}{n} \log \mu_\Phi(\mathcal{P}^{(n)}(x))\, d\eta = \int h_{\mu_\Phi}(\sigma, x)\, d\eta. \tag{3.6}$$

In particular we have $\mathcal{F}_*(\mu_\Phi, \Psi) = h_{\mu_\Phi}(\sigma)$. Our first main result is as follows.

Theorem A. *(Large deviation principle for S-M-B entropy formula) Let Φ be an almost additive sequence of continuous potentials with bounded distortion. Assume that μ_Φ is the unique equilibrium state for σ with respect to Φ and it is a Gibbs measure. Then for any $c > 0$*

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log \mu_\Phi &\left(\left\{ x \in \Sigma_A^+ : \big| -\frac{1}{n} \log \mu_\Phi(\mathcal{P}^{(n)}(x)) - h_{\mu_\Phi}(\sigma) \big| \geq c \right\} \right) \\ &\leq - \inf_{\eta \in \mathcal{M}_\sigma} \left\{ P(\Phi) - h_\eta(\sigma) - \mathcal{F}_*(\eta, \Phi) \colon \big| \int h_{\mu_\Phi}(\sigma, x)\, d\eta - h_{\mu_\Phi}(\sigma) \big| \geq c \right\}\end{aligned} \tag{UB}$$

is non-positive, and also

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \log \mu_\Phi &\left(\left\{ x \in \Sigma_A^+ : \big| -\frac{1}{n} \log \mu_\Phi(\mathcal{P}^{(n)}(x)) - h_{\mu_\Phi}(\sigma) \big| > c \right\} \right) \\ &\geq - \inf_{\eta \in \mathcal{M}_\sigma} \left\{ P(\Phi) - h_\eta(\sigma) - \mathcal{F}_*(\eta, \Phi) \colon \big| \int h_{\mu_\Phi}(\sigma, x)\, d\eta - h_{\mu_\Phi}(\sigma) \big| > c \right\}.\end{aligned} \tag{LB}$$

In addition, if the sequence $\Psi = \{-\log \mu_\Phi(\mathcal{P}^{(n)}(x))\}$ satisfies

(i) *$\frac{\psi_n}{n}$ is not uniformly convergent to a constant, or*
(ii) *$\inf_{\eta \in \mathcal{M}_\sigma} \mathcal{F}_*(\eta, \Psi) < \sup_{\eta \in \mathcal{M}_\sigma} \mathcal{F}_*(\eta, \Psi)$*

then there exists $\delta_* > 0$ so that for all $0 < \delta < \delta_*$

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_\Phi \left(\left\{ x \in \Sigma_A^+ : \big| -\frac{1}{n} \log \mu_\Phi(\mathcal{P}^{(n)}(x)) - h_{\mu_\Phi}(\sigma) \big| \geq \delta \right\} \right) < 0.$$

Proof. From [1] the unique equilibrium state μ_Φ for σ with respect to Φ is a Gibbs measure and, since σ is the shift, it is full supported. Since we assume that Φ is almost additive, it follows from Proposition 3.1 that the sequence $\Psi = \{-\log \mu_\Phi(\mathcal{P}^{(n)}(x))\}$ is continuous and also almost-additive. Since we assume the subshift of finite type σ to be topologically mixing then it satisfies the specification. Thus, taking into account relation (3.6), the first part of the theorem is now a consequence of Theorem B in [20].

For the second part of the theorem, assume $\inf_{\eta \in \mathcal{M}_\sigma} \mathcal{F}_*(\eta, \Psi) < \sup_{\eta \in \mathcal{M}_\sigma} \mathcal{F}_*(\eta, \Psi)$ and set $\delta_* = \max\{|\mathcal{F}_*(\eta, \Psi) - \mathcal{F}_*(\mu_\Phi, \Psi)| : \eta \in \mathcal{M}_\sigma\} > 0$ (the maximum does exist since $\eta \mapsto \mathcal{F}_*(\eta, \Psi)$ is continuous and the space of invariant probability measures endowed with the weak* topology is compact). If $\tilde{\eta}$ is a σ-invariant probability measure that attains the previous maximum and $0 < \delta < \delta_*$, using that $\mathcal{F}_*(\mu_\Phi, \Psi) = h_{\mu_\Phi}(\sigma)$ and μ_Φ is the unique equilibrium state for σ with respect to Φ it follows that

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log \mu_\Phi &\left(\left\{ x \in \Sigma_A^+ : \big| -\frac{1}{n} \log \mu_\Phi(\mathcal{P}^{(n)}(x)) - h_{\mu_\Phi}(\sigma) \big| \geq \delta \right\} \right) \\ &\leq -P(\Phi) + h_{\tilde{\eta}}(\sigma) + \mathcal{F}_*(\tilde{\eta}, \Phi) < 0.\end{aligned}$$

This completes the proof of the theorem. $\square$

Remark 3.3. *We observe that conditions (i) and (ii) in the theorem are equivalent by Lemma 2.2 in* [22]. *Moreover, the condition* $\inf_{\eta\in\mathcal{M}_\sigma}\mathcal{F}_*(\eta,\Psi)<\sup_{\eta\in\mathcal{M}_\sigma}\mathcal{F}_*(\eta,\Psi)$ *above is necessary to deduce the exponential convergence. In fact, otherwise it is not hard to check that* $\mathcal{F}_*(\eta,\Psi)=h_{\mu_\Phi}(\sigma)$ *for every* σ*-invariant probability measure* η.

In the wider context of Definition 2.1, where the Gibbs property is defined pointwisely and almost everywhere (the functions K_n are functions almost everywhere finite), the situation is much harder to analyze since we deal with almost everywhere sequences of pointwisely asymptotically additive potentials. Using Remark 3.2 we can weaken the hypothesis of Theorem A and deduce large deviations upper bounds for the convergence to entropy for weak Gibbs measures.

Theorem B. *Let* $\Phi=\{\varphi_n\}$ *be an asymptotically additive sequence of potentials on* Σ_A^+ *and* μ_Φ *the unique equilibrium state for* σ *with respect to* Φ. *Assume there are constants* $K_n\geq 1$ *so that* $\limsup_{n\to\infty}\frac{1}{n}\log K_n=0$ *and*

$$K_n^{-1}\leq\frac{\mu_\Phi(\mathcal{P}^{(n)}(x))}{e^{-nP_{\text{top}}(\sigma,\Phi)+\varphi_n(x)}}\leq K_n$$

for every $x\in\Sigma_A^+$ *and* $n\geq 1$. *Then the large deviations estimates (UB) and (LB) in Theorem A hold for the sequence* $\Psi=\{-\log\mu_\Phi(\mathcal{P}^{(n)}(x))\}$. *Futhermore, if* $\inf_{\eta\in\mathcal{M}_\sigma}\mathcal{F}_*(\eta,\Psi)<\sup_{\eta\in\mathcal{M}_\sigma}\mathcal{F}_*(\eta,\Psi)$ *then there exists* $\delta_*>0$ *so that for all* $0<\delta<\delta_*$

$$\limsup_{n\to\infty}\frac{1}{n}\log\mu_\Phi\left(\left\{x\in\Sigma_A^+:\big|-\frac{1}{n}\log\mu_\Phi(\mathcal{P}^{(n)}(x))-h_{\mu_\Phi}(\sigma)\big|\geq\delta\right\}\right)<0.$$

Proof. Consider $\psi_n(x)=-\log\mu_\Phi(\mathcal{P}^{(n)}(x))$ for any $n\geq 1$. Since Φ is asymptotically additive, then $\Psi=\{\psi_n\}$ is also asymptotically additive by Theorem 2.1 in [9]. The large deviations principle follows by Theorem B in [20]. Finally, the last claim in the theorem follows analogously as in the proof of Theorem A. □

In view of the previous theorem, we would expect a large deviations formulation concerning weak Gibbs measures could be given in terms of the tail of the functions $K_n(\cdot)$ that could be expressed by the following:

Conjecture: *Let* $\Phi=\{\varphi_n\}$ *be an almost additive sequence of potentials on* Σ_A^+ *and* μ_Φ *the unique equilibrium state for* σ *with respect to* Φ. *Assume that* μ_Φ *is a weak Gibbs measure for* σ *with respect to* Φ *on* $\Lambda\subset\Sigma_A^+$ *in the sense of Definition 2.1 and, that for any* $\varepsilon>0$ *there exists* $\gamma>0$ *(depending on* ε*) so that* $\mu_\Phi\big(\{x\colon K_n(x)>e^{\varepsilon n}\}\big)\leq e^{-\gamma n}$ *for all large* n. *Then, either the sequence* $\{-\frac{1}{n}\log\mu_\Phi(\mathcal{P}^{(n)}(x))\}$ *is uniformly convergent to a constant or, for every* $c>0$, *we have that*

$$\limsup_{n\to\infty}\frac{1}{n}\log\mu_\Phi\left(\left\{x\in M:\big|\frac{1}{n}\log\mu_\Phi(\mathcal{P}^{(n)}(x))+h_{\mu_\Phi}(\sigma)\big|\geq c\right\}\right)<0.$$

Clearly a finer pointwise version of Proposition 3.2 would be useful to obtain partial results in the direction of the previous conjecture. To finish this subsection, let us observe that the sequences $\psi_n=\log R_n(\cdot)$ are in general just measurable and need not necessarily have any sub-additivity property. In fact, if for μ-almost every x we have $\psi_n(x)/n$ convergent to entropy, for the dense set of periodic orbits it tends to zero. Thus our results cannot apply directly and to deal with this situation one shall need other methods.

3.2. **Errors in convergence to entropy: a computational problem.** The approximation of entropy by any of its limit characterizations is a fundamental problem from the computational viewpoint. Large deviations estimates provide a bound in probability but lack to provide an almost sure statement, which is the purpose of this section.

In this subsection, we endow the space Σ_A^+ with a distance $\tilde{d}$, let $\mathcal{P}$ be the natural generating partition which consists of cylinders of size one, and let $\sigma : \Sigma_A^+ \to \Sigma_A^+$ be the shift map. We say that $\sigma : (\Sigma_A^+, \tilde{d}) \to (\Sigma_A^+, \tilde{d})$ is an *expanding map* if it admits backward contraction, that is, if there exists $\lambda \in (0,1)$ so that for any $n \geq 1$, any $x \in X$ and any $y \in \mathcal{P}^{(n)}(x)$ it holds that

$$\tilde{d}(\sigma^j(x), \sigma^j(y)) \leq \lambda^{n-j} \tilde{d}(\sigma^n(x), \sigma^n(y)) \quad \text{for every } 1 \leq j \leq n.$$

Let μ be a σ-invariant probability measure. We say that (σ, μ) is *non-uniformly expanding* if there exists $\lambda \in (0,1)$ so that the following holds: for μ-almost every x there exists a sequence $(n_i(x))_{i \geq 1}$ satisfying

$$n_{i+1}(x)/n_i(x) \to 1 \tag{3.7}$$

such that for any $i \geq 1$ and any $y \in \mathcal{P}^{(n)}(x)$

$$\tilde{d}(\sigma^j(x), \sigma^j(y)) \leq \lambda^{n-j} \tilde{d}(\sigma^{n_i(x)}(x), \sigma^{n_i(x)}(y)) \quad \text{for every } 1 \leq j \leq n_i(x). \tag{3.8}$$

The numbers $(n_i(x))_{i \geq 1}$ are known as *hyperbolic times* for x and a sequence satisfying property (3.7) is sometimes known as non-lacunar. By definition, these hyperbolic times are locally constant: if n is a hyperbolic time for x then n is also a hyperbolic time for every $y \in \mathcal{P}^{(n)}(x)$.

Now, if (σ, μ) is non-uniformly expanding there is a μ-almost everywhere well defined *first hyperbolic time map* $n_1 = n_1(\cdot, \mathcal{P}) : (\Sigma_A^+, \tilde{d}) \to \mathbb{N}$ so that $n_1(x) = k$ if and only if $k \geq 1$ is the first integer that is a hyperbolic time for every $y \in \mathcal{P}^{(k)}(x)$. Moreover,

$$\sigma^{n_1(x)} \mid_{\mathcal{P}^{(n_1(x))}(x)} : \mathcal{P}^{(n_1(x))}(x) \to \Sigma_A^+ \tag{3.9}$$

is a bijection onto its image and an expanding map. For the distance d defined in equation (3.1) the shift is uniformly expanding and $n_1(\cdot) \equiv 1$ everywhere. In Section 5 we provide an example of a distance $\tilde{d}$ and an invariant probability measure μ under which the shift becomes a non-uniformly expanding map.

Almost sure error estimates. In what follows we address the computational problem of estimating the error in the Shannon-McMillan-Breiman limit formula for metric entropy, that is, the order of convergence of the error $r(x, n)$ in

$$h_\mu(\sigma) = h_\mu(\sigma, \mathcal{P}) = -\frac{1}{n} \log \mu(\mathcal{P}^{(n)}(x)) + r(x, n).$$

We shall focus on the case of weak Gibbs measures for non-uniformly expanding shifts to prove, roughly, that the error term $r(x, n)$ is almost everywhere bounded by the non-lacunarity $(n_{i+1}(x) - n_i(x))/n_i(x)$ of the first hyperbolic time map due to lack of uniform hyperbolicity and the term $\sqrt{\frac{\log \log n}{n}}$ due to independence of the process. To the best of our knowledge, this result is new even for uniformly expanding maps and Gibbs measures, in which case $r(x,n) = \mathcal{O}(\sqrt{\frac{\log\log n}{n}})$ almost everywhere. Let us state our main result precisely.

Theorem C. *Let φ be a Hölder continuous potential on $(\Sigma_A^+, \tilde{d})$ and assume that there exists a unique equilibrium state μ_φ for σ with respect to φ so that*

(a) *μ_φ is equivalent to a reference measure ν_φ with a Jacobian $J_{\nu_\varphi}\sigma = e^{P(\varphi)-\varphi}$, where $P(\varphi) = P_{\text{top}}(\sigma,\varphi)$,*
(b) *(σ, μ_φ) a non-uniformly expanding shift, and*
(c) *the first hyperbolic time map satisfies $n_1 \in L^{2+\delta}(\nu_\varphi)$.*

If $n_i(\cdot)$ denotes the i-th hyperbolic time map, then for μ_φ-almost every x

$$h_{\mu_\varphi}(\sigma) = -\frac{1}{n}\log \mu_\varphi(\mathcal{P}^{(n)}(x)) + r(x,n)$$

where $r(x,n) = \mathcal{O}\Big(\frac{n_{i+1}(x)-n_i(x)}{n_i(x)} + \sqrt{\frac{\log\log n}{n}}\Big)$. In the case that the subshift of finite type $\sigma : (\Sigma_A^+, \tilde{d}) \to (\Sigma_A^+, \tilde{d})$ is uniformly expanding, then $r(x,n) = \mathcal{O}\Big(\sqrt{\frac{\log\log n}{n}}\Big)$.

Proof. We first prove that under the previous conditions the measure μ_φ is a weak Gibbs measure. Let $n \geq 1$ be a hyperbolic time for x. Since $J_{\nu_\varphi}\sigma$ is Hölder continuous and there is uniform backward contraction at hyperbolic times, there exists a uniform $C > 0$ such that $C^{-1}J_{\nu_\varphi}\sigma^n(y) \leq J_{\nu_\varphi}\sigma^n(x) \leq CJ_{\nu_\varphi}\sigma^n(y)$ for any $y \in \mathcal{P}^{(n)}(x)$. Hence, using that $\sigma^n \mid_{\mathcal{P}^{(n)}(x)}$ is injective,

$$\begin{aligned} 1 =& \nu_\varphi(\sigma^n(\mathcal{P}^{(n)}(x))) = \int_{\mathcal{P}^{(n)}(x)} J_{\nu_\varphi}\sigma^n \, d\nu_\varphi \\ &\leq C\, J_{\nu_\varphi}\sigma^n(x)\, \nu_\varphi(\mathcal{P}^{(n)}(x)) = C\, \frac{\nu_\varphi(\mathcal{P}^{(n)}(x))}{e^{-nP(\varphi)+S_n\varphi(x)}} \end{aligned}$$

Since the other inequality is completely analogous this proves that there exists a uniform constant $C > 0$ so that

$$C^{-1} \leq \frac{\nu_\varphi(\mathcal{P}^{(n)}(x))}{e^{-nP(\varphi)+\varphi_n(x)}} \leq C$$

for any n hyperbolic time for $x \in \Sigma_A^+$. Since $\mu_\varphi-$almost every x has infinitely many hyperbolic times, if for any $n \geq 1$ one considers $n_i(x) \leq n < n_{i+1}(x)$ (where $n_i(x)$ denotes the ith hyperbolic time for x) then

$$\begin{aligned} \nu_\varphi(\mathcal{P}^{(n)}(x)) &\leq \nu_\varphi(\mathcal{P}^{(n_i(x))}(x)) \leq C\, e^{-n_i(x)P(\varphi)+S_{n_i(x)}\varphi(x)} \\ &\leq C_n(x)\, e^{-nP(\varphi)+S_n\varphi(x)} \end{aligned}$$

with $C_n(x) = C\, e^{[n-n_i(x)]P(\varphi)-S_{[n-n_i(x)]}\varphi(\sigma^{n_i(x)}(x))} \leq C\, e^{(|P(\varphi)|+\|\varphi\|)(n_{i+1}(x)-n_i(x))}$. The other inequality follows analogously. Furthermore,

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n}\log C_n(x) \quad &\leq \quad \limsup_{n\to\infty} \frac{n_{i+1}(x)-n_i(x)}{n}(|P(\varphi)|+\|\varphi\|) \\ &\leq \quad \limsup_{i\to\infty} \frac{n_{i+1}(x)-n_i(x)}{n_i(x)}(|P(\varphi)|+\|\varphi\|) = 0 \end{aligned}$$

for μ_φ-almost every x. Using that μ_φ and ν_φ are equivalent probability measures, hence $\frac{d\mu_\varphi}{d\nu_\varphi}$ is bounded by a uniform constant K, for μ_φ-almost every x there exists a sequence of positive constants $\{K_n(x)\}_{n\geq 1}$ (depending only on x)

$$K_n(x) = C\,K\; e^{(|P(\varphi)|+\|\varphi\|)(n_{i+1}(x)-n_i(x))} \tag{3.10}$$

satisfying $\lim_{n\to\infty} \frac{1}{n}\log K_n(x) = 0$ and so that

$$K_n(x)^{-1} \leq \frac{\mu_\varphi(\mathcal{P}^{(n)}(x))}{e^{-nP(\varphi)+S_n\varphi(x)}} \leq K_n(x)$$

for every $n \geq 1$. Since μ_φ is an equilibrium state for σ with respect to φ, then $P(\varphi) = h_{\mu_\varphi}(\sigma) + \int \varphi\, d\mu_\varphi$ and so

$$\begin{aligned} S_n\varphi(x) - n\int \varphi\, d\mu_\varphi - \log K_n(x) &\leq n h_{\mu_\varphi}(\sigma) + \log \mu_\varphi(\mathcal{P}^{(n)}(x)) \\ &\leq S_n\varphi(x) - n\int \varphi\, d\mu_\varphi + \log K_n(x) \end{aligned}$$

for μ_φ-almost every $x \in \Sigma_A^+$ and every $n \geq 1$. In particular, the error term $r(x,n) = h_{\mu_\varphi}(\sigma) + \frac{1}{n}\log \mu_\varphi(\mathcal{P}^{(n)}(x))$ verifies

$$\frac{1}{n} S_n\varphi(x) - \int \varphi\, d\mu_\varphi - \frac{1}{n}\log K_n(x) \leq r(x,n) \leq \frac{1}{n} S_n\varphi(x) - \int \varphi\, d\mu_\varphi + \frac{1}{n}\log K_n(x). \tag{3.11}$$

Let $R(\cdot) = n_1(\cdot)$ be the first hyperbolic time map given by equations (3.8) and (3.9). Then (σ, μ_φ) admits an induced transformation $F = \sigma^{R(\cdot)} \colon \Sigma_A^+ \to \Sigma_A^+$ that is Gibbs-Markov: (i) there exists a partition $\mathcal{D}$ (mod 0) of Σ_A^+ so that $F\mid_D \colon D \to F(D) \subset \Sigma_A^+$ is a bijection and is uniformly expanding for each $D \in \mathcal{D}$; (ii) $F(D) \in \mathcal{D}$ for every $D \in \mathcal{D}$ (big images), and (iii) $\log J_{\nu_\varphi}\sigma$ is Hölder continuous with respect to the metric $\tilde{d}$ (bounded distortion) (c.f. [14, Section 2]). By assumption on the first hyperbolic time map, one has $R \in L^{2+\delta}(\nu_\varphi)$. By the almost sure invariance principle (and consequent law of iterated logarithm c.f. [16]) proved by Melbourne and Nicol [14] for non-uniformly expanding maps, the mean zero and Hölder continuous observable $\tilde{\varphi} = \varphi - \int \varphi\, d\mu_\varphi$ verifies

$$\limsup_{n\to\infty} \frac{1}{\sqrt{n\log\log n}} \Big| \sum_{j=0}^{n-1} \tilde{\varphi}(\sigma^j(x)) \Big| = \sqrt{2}.$$

for μ_φ-almost every x (we just observe that the starting point in [14] is a Lebesgue space (Λ, m) which in our setting can be taken $\Lambda = \Sigma_A^+$ and $m = \nu_\varphi$ while the transfer operator estimates hold in our context with $g = \log J_{\nu_\varphi}\sigma$).

On the one hand, for μ_φ-a.e. x there exists $n(x) \geq 1$ so that

$$\frac{1}{n}\Big| \sum_{j=0}^{n-1} \tilde{\varphi}(f^j(x)) \Big| \leq 2\sqrt{\frac{\log\log n}{n}}$$

for every $n \geq n(x)$. On the other hand, notice that if $n_i(x) \leq n < n_{i+1}(x)$ are consecutive hyperbolic times for x, then for sufficiently large n it follows from (3.10) and $\frac{n_{i+1}(x)}{n_i(x)} \to 1$ that

$$\frac{1}{n}\log K_n(x) \lesssim \frac{n_{i+1}(x) - n_i(x)}{n_i(x)}.$$

Putting altogether in equation (3.11), this yields the desired result. $\square$

Remark 3.4. *The reference measure ν_φ considered in the previous theorem is sometimes called conformal measure. In the case of smooth non-uniformly expanding maps on manifolds, the reference measure ν_φ usually coincides with Leb, in which case the Jacobian is $J_{\nu_\varphi} f = \log|\det Df|$. In the case of subshifts of finite type, conformal measures arise as fixed points of the dual of the Ruelle-Perron-Frobenius operator (see e.g.* [4]*).*

Remark 3.5. *By the definition of hyperbolic times in* (3.9)*, these are not only constant on cylinders of the Markov partition but are also additive, that is, $n_{i+1}(x) = n_i(x) + n_1(\sigma^{n_i(x)}(x))$ for every $i \geq 1$ and μ_φ-almost every x. Hence, the*

$$\frac{n_{i+1}(x) - n_i(x)}{n_i(x)} \leq \frac{n_1(\sigma^{n_i(x)}(x))}{n_1(x)}$$

can be expressed in terms of the first hyperbolic time map.

Remark 3.6. *The next table illustrates the speed of convergence to zero of the term $r_n := \sqrt{\frac{\log\log n}{n}}$ as n tends to infinity (here* log *are computed in base e):*

n	r_n
10^2	$\approx 0.1235791093109147$
10^3	$\approx 0.0439618554421451$
10^4	$\approx 0.0149007610757567$
10^5	≈ 0.00494314677
10^5	$\approx 0.0049431471328315$
10^6	$\approx 0.0016204295462858$
10^7	$\approx 0.000527251609225$
10^8	$\approx 0.0001706890150809$
10^9	$\approx 0.0000550568526397$
10^{10}	$\approx 0.000017710498407$
10^{20}	$\approx 0.0000000000195697846668$
10^{30}	$\approx 0.00000000000000020579674018093947$

4. Local pressure and topological formulation of large deviations

Throughout this section, let $f : X \to X$ be a continuous transformation on a compact metric space X. We introduce a capacity pressure for a family of sets for general continuous dynamical systems.

Let $\Phi = \{\varphi_n\}$ be a sequence of continuous potentials on X and let $\{Z_n\}_n$ be a family of subsets of X. Given a finite open cover $\mathcal{U}$ of X and $n \geq 1$, set $\mathcal{U}^{(n)} = \bigvee_{i=0}^{n-1} f^{-i}\mathcal{U}$. Consider

$$\overline{CP}(f, \Phi, \{Z_n\}, \mathcal{U}) = \limsup_{n\to\infty} \frac{1}{n} \log \inf_{\mathcal{G}_n} \Big\{ \sum_{U \in \mathcal{G}_n} e^{\varphi_n(U)} \Big\} \tag{4.1}$$

where the infimum is taken over all subfamilies $\mathcal{G}_n$ of $\mathcal{U}^{(n)}$ that cover Z_n and $\varphi_n(U) = \sup_{x\in U} \varphi_n(x)$. Finally consider the *upper capacity pressure of* $\{Z_n\}_n$ by

$$\overline{CP}(f, \Phi, \{Z_n\}) = \liminf_{|\mathcal{U}|\to 0} \overline{CP}(f, \Phi, \{Z_n\}, \mathcal{U}) \tag{4.2}$$

where $|\mathcal{U}|$ denotes the diameter of the open cover $\mathcal{U}$.

Remark 4.1. *It follows from the definition of upper capacity pressure together with the results in* [5] *that:*

- (i) *if $Z_n \subset \tilde{Z}_n$ for all n large then $\overline{CP}(f, \Phi, \{Z_n\}) \leq \overline{CP}(f, \Phi, \{\tilde{Z}_n\})$;*
- (ii) *if $Z_n = X$ for all n large then $\overline{CP}(f, \Phi, X) \geq P_{\text{top}}(f, \Phi)$ for every subadditive sequence of continuous potentials $\Phi = \{\varphi_n\}$;*
- (iii) *for any family $\{Z_n\}$ of subsets of X, if either*

a. Φ *is asymptotically additive, or*
b. Φ *is subadditive and has weak bounded distortion condition, or*
c. *the topological entropy of f is finite and the metric entropy function $\mu \mapsto h_\mu(f)$ is upper semicontinuous*

then $\overline{CP}(f, \Phi, \{Z_n\}) \leq \overline{CP}(f, \Phi, X) = P_{\text{top}}(f, \Phi)$.

The first result is clear from the definitions. See Lemma 4.3 in [5] *for the proof of the second result. The first inequality in (iii) is clear from the definitions. Finally, the equality* $\overline{CP}(f, \Phi, X) = P_{\text{top}}(f, \Phi)$ *follows from Lemma 2.1 in* [22] *and Proposition 4.7 in* [5] *in case a, and follows from Proposition 4.7 and Proposition 4.4 in* [5], *respectively, in cases b and c.*

We now describe an equivalent definition of upper capacity pressure of a sequence of subsets $\{Z_n\}_n$. Given a sequence $\{Z_n\}$ of subsets of X, for each $N \geq 1$ set

$$R(\Phi, \{Z_n\}, N, \delta) = \inf \Big\{ \sum_i \exp\big(\sup_{y \in B_N(x_i, \delta)} \varphi_N(y) \big) : \bigcup_i B_N(x_i, \delta) \supset Z_N \Big\}. \quad (4.3)$$

Using standard arguments one can prove that

$$\overline{CP}(f, \Phi, \{Z_n\}) = \liminf_{\delta \to 0} \limsup_{N \to \infty} \frac{1}{N} \log R(\Phi, \{Z_n\}, N, \delta).$$

Let us mention that if $Z_n = Z$ for each n, we simply write the upper capacity pressure of $\{Z_n\}_n$ as $\overline{CP}(f, \Phi, Z)$. Moreover, in the case of the potentials $\varphi_n \equiv 0$ and sets $Z_n = B_n(x, r)$ for some fixed $x \in M$ and $r > 0$ then

$$\overline{CP}(f, \Phi, \{Z_n\}) = \lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \log \inf \sharp \mathcal{G}_n$$

where the infimum above is taken over all subfamilies $\mathcal{G}_n$ of (n, δ)-dynamical balls necessary to cover $B_n(x, r)$. Thus, in this case the upper capacity pressure coincides with the r-local entropy at x, usually denoted by $h_{\text{loc}}(f, x, r)$. The *local entropy* at point x is defined by $h_{\text{loc}}(f, x) := \lim_{r \to 0} h_{\text{loc}}(f, x, r)$

Given an f-invariant ergodic measure μ, in our context the study of large deviations concerns the velocity at which the measure of the deviation sets

$$\Gamma_{c,n} = \Big\{ x \in X \colon \big| \frac{1}{n} \psi_n(x) - \mathcal{F}_*(\mu, \Psi) \big| \geq c \Big\} \quad (4.4)$$

decrease to zero for some sequence $\Psi = \{\psi_n\}$ provided the convergence in Kingman's subadditive ergodic theorem holds. Our next result relates the (probabilistic) large deviations results for Gibbs measures with the (topological) upper capacity pressure asymptotics of the deviation sets. More precisely,

Theorem D. *Let $\Phi = \{\varphi_n\}$ be a sequence of continuous potentials. Assume that μ_Φ is a weak Gibbs measure for f with respect to Φ: for any $0 < \varepsilon < \varepsilon_0$ there are $K_n = K_n(\varepsilon) \geq 1$ satisfying $\limsup_{n \to \infty} \frac{1}{n} \log K_n(\varepsilon) = 0$ and so that*

$$K_n^{-1} \leq \frac{\mu_\Phi(B_n(x, \varepsilon))}{e^{-nP_{\text{top}}(f, \Phi) + \varphi_n(x)}} \leq K_n$$

for every $x \in X$, $0 < \varepsilon < \varepsilon_0$ and $n \geq 1$. Given $c > 0$, if the deviation sets $\Gamma_{c,n}$ are defined by equation (4.4) then the following large deviations upper bound holds:

$$\limsup_{n \to \infty} \frac{1}{n} \log \mu_\Phi(\Gamma_{c,n}) \leq -P_{\text{top}}(f, \Phi) + \overline{CP}(f, \Phi, \{\Gamma_{c,n}\}).$$

In particular,

(a) *if Φ is asymptotically additive then* $\limsup_{n\to\infty} \frac{1}{n} \log \mu_\Phi(\Gamma_{c,n}) \le 0$.

(b) *if* $\overline{CP}(f,\Phi,\{\Gamma_{c,n}\}) < P_{\text{top}}(f,\Phi)$ *then* $\mu_\Phi(\Gamma_{c,n})$ *decreases exponentially fast.*

Conversely, if Φ satisfies the weak bounded distortion property and $\mu_\Phi(\Gamma_{c,n})$ decreases exponentially fast then $\overline{CP}(f,\Phi,\{\Gamma_{c+\gamma,n}\}) < P_{\text{top}}(f,\Phi)$ *for any* $\gamma > 0$.

Proof. Let $c > 0$ be fixed. For each positive integer n and $0 < \delta < \varepsilon_0$, let $\{B_n(x_i,\delta)\}_{i\in\mathcal{I}}$ be a cover of $\Gamma_{c,n}$. Using the Gibbs property of the measure μ_Φ, we have

$$\mu_\Phi(\Gamma_{c,n}) \le \sum_{i\in\mathcal{I}} \mu_\Phi(B_n(x_i,\delta)) \le K_n e^{-nP_{\text{top}}(f,\Phi)} \sum_{i\in\mathcal{I}} \exp\Big(\sup_{y\in B_n(x_i,\delta)} \varphi_n(y)\Big). \quad (4.5)$$

Since the cover by balls was chosen arbitrary then

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_\Phi(\Gamma_{c,n}) \le -P_{\text{top}}(f,\Phi) + \limsup_{n\to\infty} \frac{1}{n} \log R(\Phi,\{\Gamma_{c,n}\},n,\delta).$$

Letting δ tend to zero we deduce that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_\Phi(\Gamma_{c,n}) \le -P_{\text{top}}(f,\Phi) + \overline{CP}(f,\Phi,\{\Gamma_{c,n}\}).$$

Item (a) in the theorem follows from the previous estimate and item (iii) in Remark 4.1. Item (b) follows immediately from the last estimate above.

To complete the proof of the theorem, let $\Phi = \{\varphi_n\}$ be a family of continuous potentials satisfying the weak bounded distortion property, let $c > 0$ be such that $\mu_\Phi(\Gamma_{c,n})$ decreases exponentially fast and let $\gamma > 0$ be arbitrary. There exists $\delta_0 > 0$ and $n_0 \ge 1$ so that $\gamma_n(\Phi,\delta) \le \gamma\, n$ for every $n \ge n_0$ and $0 < \delta < \delta_0$.

If $(x_i)_{i\in\mathcal{I}}$ is a maximal (n,δ)-separated subset of $\Gamma_{c+\gamma,n}$ then the elements of $\{B_n(x_i,\delta/2)\}_{i\in\mathcal{I}}$ are pairwise disjoint and $B_n(x_i,\delta) \subset \Gamma_{c,n}$. This yields that

$$\begin{aligned}
\mu_\Phi(\Gamma_{c,n}) &\ge \sum_{i\in\mathcal{I}} \mu_\Phi(B_n(x_i,\delta/2)) \ge K_n^{-1} e^{-nP_{\text{top}}(f,\Phi)} \sum_{i\in\mathcal{I}} e^{\varphi_n(x_i)} \\
&\ge K_n^{-1} e^{-nP_{\text{top}}(f,\Phi)} e^{-\gamma_n(\Phi,\delta)} \sum_{i\in\mathcal{I}} \exp\Big(\sup_{y\in B_n(x_i,\delta)} \varphi_n(y)\Big) \\
&\ge K_n^{-1} e^{-nP_{\text{top}}(f,\Phi)} e^{-\gamma_n(\Phi,\delta)} R(\Phi,\{\Gamma_{c+\gamma,n}\},n,\delta),
\end{aligned}$$

and consequently

$$\begin{aligned}
\limsup_{n\to\infty} \frac{1}{n} \log \mu_\Phi(\Gamma_{c,n}) \ge -P_{\text{top}}(f,\Phi) &- \limsup_{n\to\infty} \frac{\gamma_n(\Phi,\delta)}{n} \\
&+ \limsup_{n\to\infty} \frac{1}{n} \log R(\Phi,\{\Gamma_{c+\gamma,n}\},n,\delta).
\end{aligned}$$

Letting $\delta \to 0$, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_\Phi(\Gamma_{c,n}) \ge -P_{\text{top}}(f,\Phi) + \overline{CP}(f,\Phi,\{\Gamma_{c+\gamma,n}\}).$$

By assumption the measure μ_Φ of the sets $\Gamma_{c,n}$ decreases exponentially fast and, consequently, $\overline{CP}(f,\Phi,\{\Gamma_{c+\gamma,n}\}) < P_{\text{top}}(f,\Phi)$. This completes the proof of the theorem. □

We remark that for hyperbolic dynamics f, sequences of almost-additive continuous potentials $\Psi = \{\psi_n\}$ admit a weak Gibbs measure. In what follows we give geometrical and topological properties for weak Gibbs measures. The first one is that positive measure sets for weak Gibbs measures are topologically large, namely that carry full capacity pressure.

Corollary A. *Let $\Phi = \{\varphi_n\}$ be an almost additive sequence of continuous potentials on X. Assume μ_Φ is a weak Gibbs measure for (f, Φ) as in Theorem D. If $A \subset X$ is a positive μ_Φ-measure set then*

$$\overline{CP}(f, \Phi, A) = P_{\text{top}}(f, \Phi).$$

In particular, if μ_0 is a maximal entropy measure and $\mu_0(A) > 0$ then A has upper pressure capacity equal to $h_{\text{top}}(f)$.

Proof. Fix a set $A \subset X$ of positive $\mu_\Phi-$measure. Since Φ is almost additive, it follows from item (iii) in Remark 4.1 that $\overline{CP}(f, \Phi, A) \le \overline{CP}(f, \Phi, X) = P_{\text{top}}(f, \Phi)$. Taking $\Gamma_n = A$ for every $n \ge 1$, by the same arguments as in the proof of Theorem D it yields that

$$0 = \limsup_{n\to\infty} \frac{1}{n} \log \mu_\Phi(A) \le -P_{\text{top}}(f, \Phi) + \overline{CP}(f, \Phi, A).$$

Consequently, we have $P_{\text{top}}(f, \Phi) = \overline{CP}(f, \Phi, A)$. Since the second assertion is immediate from the first one, this finishes the proof of the corollary. □

The second property, on the geometry of weak Gibbs measures, is on rough terms that Gibbs measures exhibit some exact dimensionality with respect to dynamical balls, which allows one to prove that the r-local entropy is zero almost everywhere. More precisely,

Corollary B. *Let Φ be an asymptotically additive sequence of continuous functions. If μ_Φ is a weak Gibbs measure for f with respect to Φ as in Theorem D then $h_{loc}(f, x, r) = 0$ for μ_Φ-almost every $x \in X$ and every small $r > 0$.*

Proof. Given $0 < \varepsilon < r$ small, $x \in X$ and $n \ge 1$, let $\{B_n(x_i, \varepsilon)\}_{i\in\mathcal{I}}$ be a finite minimal cover of $B_n(x, r)$ ($\mathcal{I}$ depends on x, r, n). Observe that

$$B_n(x, r) \subset \bigcup_{i\in\mathcal{I}} B_n(x_i, \varepsilon) \subset B_n(x, 2r)$$

and the collection $\{B_n(x_i, \frac{\varepsilon}{2})\}_{i\in\mathcal{I}}$ is pairwise disjoint. Since $x_i \in B_n(x, 2r)$ it follows from the Gibbs property that

$$K_n(\varepsilon/2)\ e^{-nP(\Phi)+\varphi_n(x)}\ e^{-\gamma_n(\Phi,2r)} \le \mu_\Phi(B_n(x_i, \frac{\varepsilon}{2})) \le K_n(\varepsilon/2)\ e^{-nP(\Phi)+\varphi_n(x)}\ e^{\gamma_n(\Phi,2r)}.$$

This, together with the fact that

$$\sum_{i\in\mathcal{I}} \mu_\Phi(B_n(x_i, \varepsilon/2)) \le \mu_\Phi(B_n(x, 2r)) \le K_n(2r)\ e^{-nP(\Phi)+\varphi_n(x)}$$

implies that the minimum cardinality $N_n(x, r, \varepsilon)$ of dynamical balls of radius ε necessary to cover $B_n(x, r)$ is bounded from above by $K_n(2r)K_n(\varepsilon/2)^{-1}\ e^{\gamma_n(\Phi,2r)}$ for every $n \ge 1$. In consequence,

$$h_{\text{loc}}(f, x) = \lim_{r\to 0} \limsup_{n\to\infty} \frac{1}{n} \log N_n(x, r, \varepsilon) = 0$$

which completes the proof of the corollary. □

5. Examples and applications

In this section we give some examples that illustrate the concepts involved and where our results apply. First we estimate the rate of convergence for all Lyapunov exponents associated to locally constant linear cocycles.

Example 5.1. *(Convergence of all Lyapunov exponents) Let $\sigma : \Sigma \to \Sigma$ be the shift map on the space $\Sigma = \{1, \dots, \ell\}^{\mathbb{N}}$ endowed with the distance $d(x, y) = 2^{-n}$ where $x = (x_j)_j$, $y = (y_j)_j$ and $n = \min\{j \geq 0 : x_j \neq y_j\}$. Let μ_φ be the unique Gibbs equilibrium state for σ with respect to a Hölder continuous potential φ.*

Let $A : \Sigma \to SL(\ell, \mathbb{R})$ be a locally constant linear cocycle, that is $A \mid_{[i]} = M_i$ for every $1 \leq i \leq \ell$. Given $\iota = (i_1, \dots, i_n) \in \{1, \dots, \ell\}^n$, consider the matrix

$$A^{(n)}(\iota) := M_{i_n} \dots M_{i_2} M_{i_1}.$$

By ergodicity of μ_φ and Oseledets theorem, there exists $1 \leq k \leq \ell$ and for μ_ϕ-a.e. ι there exists a cocycle invariant decomposition $\mathbb{R}^\ell = E^1_\iota \oplus E^2_\iota \oplus \dots \oplus E^k_\iota$ and real numbers $\lambda_1 > \lambda_2 > \dots > \lambda_k$ so that

$$\lambda_i := \lambda_i(A, \mu_\varphi) = \lim_{n\to\infty} \frac{1}{n} \log \|A^{(n)}(\iota)\, v_i\|$$

for every $v_i \in E^i_\iota \setminus \{0\}$. Moreover, for μ_ϕ-a.e. ι

$$\lim_{n\to\infty} \frac{1}{n} \log \|\Lambda^j A^{(n)}(\iota)\| = \sum_{i=1}^{j} \lambda_i(A, \mu_\varphi)$$

where Λ^j denotes the jth exterior power. Given $1 \leq j \leq \ell$, the sequence of continuous observables $\Psi^j = \{\log \|\Lambda^j A^{(n)}(\cdot)\|\}$ is sub-additive and locally constant. Hence, it follows from [20] the exponential convergence to the sum of all j-larger Lyapunov exponents. The authors are grateful to I. Morris for pointing out this fact.

Our next result indicates how Theorem A can be used to estimate the velocity of convergence for entropy formulas in the uniformly expanding context.

Example 5.2. *Consider the circle $S^1 = [0, 1]/\sim$, where the equivalence relation $\sim$ means that the extremal points in the interval are identified. Given an expanding map f on S^1 and an open interval $I \subset [0, 1]$ so that $f \mid_I$ is injective and $f(I) = (0, 1)$, notice that $\Lambda = \bigcap_{n\geq 0} f^{-n}(S^1 \setminus I)$ is an f-invariant Cantor set and $f \mid_\Lambda$ is expanding. Then the system (Λ, f) is topologically conjugate to the one-sided full shift on two symbols (Σ_2^+, σ) via an homeomorphism $\pi : \Sigma_2 \to \Lambda$. Let $\mathcal{P}$ be the natural generating partition of Σ_2^+ consisting of cylinders of length one. Given an Hölder continuous function $\varphi : \Lambda \to \mathbb{R}$, let μ_φ be the unique Gibbs equilibrium state for $f \mid_\Lambda$ with respect to φ.*

By construction, the family $\Psi = \{\psi_n\}_n$ with $\psi_n(\omega) = -\log \mu_\varphi(\pi(\mathcal{P}^{(n)}(\omega)))$ is locally constant and almost additive. Thus

$$-\frac{1}{n} \log K \leq -\frac{1}{n} \psi_n(\omega) + P_{\text{top}}(f, \varphi) - \frac{1}{n} S_n \varphi(\pi(\omega)) \leq \frac{1}{n} \log K$$

for every $n \geq 1$ and $\omega \in \Sigma_2^+$. This implies that φ is cohomogous to a constant if and only if $\frac{\psi_n}{n}$ is uniformly convergent to a constant. Thus, using Remark 3.3, if φ is not cohomologous to a constant then $\inf_\eta \mathcal{F}_*(\eta, \Psi) < \sup_\eta \mathcal{F}_*(\eta, \Psi)$ and it follows from Theorem A that there exists $\delta_* > 0$ so that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_\varphi\Big(\Big\{x \in \Lambda : \big| -\frac{1}{n} \log \mu_\varphi(\pi(\mathcal{P}^{(n)}(\omega))) - h_{\mu_\varphi}(f)\big| \geq \delta, \pi(\omega) = x\Big\}\Big) < 0$$

for all $0 < \delta < \delta_*$.

The following example shows how Theorem B can be used for an expanding Markov maps with indifferent fixed points and non-additive sequence of continuous potentials.

Example 5.3. *Let* $I = [0,1]$ *and* $\{I_i\}_{i=1}^k$ *a family of disjoint closed intervals of* I. *Assume that* $f : \bigcup_i^k I_i \to I$ *is an expanding Markov map, that is:*

(1) *the map* f *is piecewise* C^1;
(2) f *is Markov and it can be coded by a topologically mixing sub-shift of finite on the alphabet* $\{1, \cdots, k\}$;
(3) *for every* $x \in \bigcup_i^k I_i$ *we have that* $|f'(x)| \geq 1$ *and there exists at most finitely many points for which* $|f'(x)| = 1$.

Consider the associated repeller $\Lambda := \{x \in \bigcup_i^k I_i : \ f^n(x)$ *is well-defined for every* $n \in \mathbb{N}\}$. *Then there exists a topologically mixing subshift of finite type* (Σ_A^+, σ) *defined on the alphabet* $\{1, \cdots, k\}$ *which is conjugate with the system* (Λ, f) *with a homeomorphism* $\pi : \Sigma_A^+ \to \Lambda$. *For each cylinder* $C_n(\omega) \subset \Sigma_A^+$, *let* $I_n(\omega) = \pi(C_n(\omega))$ *denote the cylinder of length* n *in* $[0,1]$ *containing the point* $\pi(\omega)$. *Let* μ_ϕ *be a weak Gibbs measure corresponding to the continuous potential* $\phi : \Lambda \to \mathbb{R}$, *i.e., there are positive constants* K_n *satisfying* $\lim_{n\to\infty} \frac{1}{n} \log K_n = 0$ *so that*

$$K_n^{-1} \leq \frac{\mu_\phi(I_n(\omega))}{\exp(-nP_{\text{top}}(f, \phi) + S_n\phi(x))} \leq K_n, \ \forall n \geq 1, \ \forall \omega \in \Sigma_A^+$$

where $x = \pi(\omega)$. *Consider* $\bar{\mu} = \mu_\phi \circ \pi$, *then*

$$K_n^{-1} \leq \frac{\bar{\mu}(C_n(\omega))}{\exp(-nP_{\text{top}}(\sigma, \phi\circ\pi) + S_n\phi(\pi(\omega)))} \leq K_n, \ \forall n \geq 1, \ \forall \omega \in \Sigma_A^+.$$

By Remark 3.2, $\Psi = \{-\log \bar{\mu}(\mathcal{P}^{(n)}(\omega))\}$ *is an asymptotically additive sequence of continuous functions, here* $\mathcal{P}$ *is the natural generating partition of* Σ_A^+ *consisting of cylinders of length one. If* $\inf_\eta \mathcal{F}_*(\eta, \Psi) < \sup_\eta \mathcal{F}_*(\eta, \Psi)$, *it follows from Theorem B that there exists* $\delta_* > 0$ *so that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \bar{\mu}\Big(\Big\{\omega \in \Sigma_A^+ : \big| -\frac{1}{n}\log \bar{\mu}(\mathcal{P}^{(n)}(\omega)) - h_{\bar{\mu}}(\sigma)\big| \geq \delta\Big\}\Big) < 0$$

for all $0 < \delta < \delta_*$. *Consequently, that there exists* $\delta_* > 0$ *so that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_\phi\Big(\Big\{x \in \Lambda : \big| -\frac{1}{n}\log \mu_\phi(I_n(\omega) - h_{\mu_\phi}(f)\big| \geq \delta, \pi(\omega) = x\Big\}\Big) < 0$$

for all $0 < \delta < \delta_*$.

It follows from Lemma 8 in [10] *that the pointwise dimension of* μ_ϕ *can be almost everywhere computed by the quotient of two asymptotically additive sequences, namely*

$$d_{\mu_\phi}(x) = \lim_{r\to 0} \frac{\log \mu_\phi((x-r, x+r))}{\log r} = \lim_{n\to\infty} \frac{\log \bar{\mu}(C_n(\omega))}{\log \operatorname{diam}(I_n(\omega))} \tag{5.1}$$

where $C_n(\omega)$ *is the cylinder of length* n *that contains* $\pi^{-1}(x)$ *(c.f. Lemmas 3.1 and 3.2 in* [9]*).*

In the case that the pointwise dimension is a constant almost everywhere (e.g. if μ_ϕ is ergodic then $d_{\mu_\phi}(x) = \frac{h_{\mu_\phi}(f)}{\int \log|f'|\,d\mu_\phi}$ for μ_ϕ-almost every x). Some easy computations guarantee that

$$\Big\{\omega \in \Sigma_A^+ : \Big|\frac{\log \mu_\phi(I_n(\omega))}{\log \operatorname{diam}(I_n(\omega))} - \frac{h_{\mu_\phi}(f)}{\int \log|f'|d\mu_\phi}\Big| > \delta\Big\}$$

is contained in the union of the sets

$$\begin{aligned}(I) &= \Big\{\omega \in \Sigma_A^+ : |\frac{1}{n}\log\operatorname{diam}(I_n(\omega)) - \int \log|f'|d\mu_\phi| > \delta^2\Big\}\\ &= \{\omega \in \Sigma_A^+ : |\frac{1}{n}\varphi_n(\omega) - \mathcal{F}_*(\mu_\phi,\Phi)| > \delta^2\}\end{aligned}$$

and

$$\begin{aligned}(II) = \Big\{\omega \in \Sigma_A^+ :& |\frac{1}{n}\log\operatorname{diam}(I_n(\omega)) - \int \log|f'|d\mu_\phi| \le \delta^2 \textit{ and}\\ &\Big|\frac{\frac{1}{n}\log\mu_\phi(I_n(\omega))}{\frac{1}{n}\log\operatorname{diam}(I_n(\omega))} - \frac{h_{\mu_\phi}(f)}{\int\log|f'|d\mu_\phi}\Big| > \delta\Big\}\end{aligned}$$

$$= \Big\{\omega \in \Sigma_A^+ : |\frac{1}{n}\varphi_n(\omega) - \mathcal{F}_*(\mu_\phi,\Phi)| \le \delta^2 \textit{ and} \Big|\frac{\psi_n}{\varphi_n} - \frac{\mathcal{F}_*(\mu_\phi,\Psi)}{\mathcal{F}_*(\mu_\phi,\Phi)}\Big| > \delta\Big\}$$

for $\varphi_n(\omega) = \log\operatorname{diam}(I_n(\omega))$ and $\psi_n(\omega) = \log\mu_\phi(I_n(\omega))$. These are continuous observables and the families $\Phi = \{\varphi_n\}$ and $\Psi = \{\psi_n\}$ are asymptotically additive. For points in (II) we have $|\frac{1}{n}\varphi_n(\omega) - \mathcal{F}_*(\mu_\phi,\Phi)| \le \delta^2$ and, consequently,

$$\begin{aligned}\delta < \Big|\frac{\psi_n}{\varphi_n} - \frac{\mathcal{F}_*(\mu_\phi,\Psi)}{\mathcal{F}_*(\mu_\phi,\Phi)}\Big| &= \Big|\frac{\frac{1}{n}\psi_n}{\frac{1}{n}\varphi_n} - \frac{\mathcal{F}_*(\mu_\phi,\Psi)}{\mathcal{F}_*(\mu_\phi,\Phi)}\Big|\\ &= \Big|\frac{\big[\frac{1}{n}\psi_n - \mathcal{F}_*(\mu_\phi,\Psi)\big] + \frac{\mathcal{F}_*(\mu_\phi,\Psi)}{\mathcal{F}_*(\mu_\phi,\Phi)}\big[\mathcal{F}_*(\mu_\phi,\Phi) - \frac{1}{n}\varphi_n\big]}{\big[\frac{1}{n}\varphi_n - \mathcal{F}_*(\mu_\phi,\Phi)\big] + \mathcal{F}_*(\mu_\phi,\Phi)}\Big|\\ &\le \frac{\big|\frac{1}{n}\psi_n - \mathcal{F}_*(\mu_\phi,\Psi)\big| + \delta^2\Big|\frac{\mathcal{F}_*(\mu_\phi,\Psi)}{\mathcal{F}_*(\mu_\phi,\Phi)}\Big|}{\big|\mathcal{F}_*(\mu_\phi,\Phi)\big| - \delta^2}\end{aligned}$$

Thus, if $\delta > 0$ is chosen small then

$$(II) \subseteq \Big\{\omega \in \Sigma_A^+ : |-\frac{1}{n}\log\mu_\phi(I_n(\omega)) - h_{\mu_\phi}(f)| > \delta\frac{|\mathcal{F}_*(\mu_\phi,\Phi)|}{2}\Big\}.$$

Both sequences Φ, Ψ are asymptotically additive, our results imply exponential large deviations rate for the convergence of the local dimension given by equation (5.1).

In what follows we provide a simple example that illustrates how to construct a metric with respect to which the shift is non-uniformly expanding that meets the assumptions of Theorem C.

Example 5.4. *Consider the local diffeomorphism f on the compact metric space $X = [0,1/3] \cup [2/3,1]$ where*

$$f(x) = \begin{cases} x[1 + 2.(3x)^\alpha] & \textit{if } 0 \le x \le \frac{1}{3}\\ 3x - 2 & \textit{if } \frac{2}{3} \le x \le 1,\end{cases}$$

where $\alpha \in (0,1)$. The set $\Lambda = \bigcap_{n\ge 0} f^{-n}(X)$ is an f-invariant Cantor set on which $|f'(0)| = 1$ and $|f'(x)| > 1$ for every $x \in \Lambda \setminus \{0\}$ (hence $f\mid_\Lambda$ presents *an intermitency phenomenon). For any Hölder continuous potential ϕ close to constant there exists a unique equilibrium state μ_ϕ that is equivalent to a weak Gibbs measure ν_ϕ whose first hyperbolic time map has exponential tail (see* [18] *for more details). Let $\Sigma_2^+ = \{0,1\}^{\mathbb{N}}$ denote the full shift space on two symbols and consider the itinerary map $\iota : \Lambda \to \Sigma_2^+$ given by $\iota(x) = (a_n)_{n\geq 1}$ with $a_n = 0$ if $f^n(x) \in [0, \frac{1}{3}]$ and $a_n = 1$ whenever $f^n(x) \in [\frac{2}{3}, 1]$. Clearly ι is a bijection and defines a distance $\tilde{d}$ on $\Sigma_2^+ = \{0,1\}^{\mathbb{N}}$ by*

$$\tilde{d}((a_n)_n, (b_n)_n) := |\iota^{-1}((a_n)_n) - \iota^{-1}((b_n)_n)|.$$

Then $\iota : (\Lambda, |\cdot|) \to (\Sigma_2^+, \tilde{d})$ is a isometry and $\sigma : (\Sigma_2^+, \tilde{d}) \to (\Sigma_2^+, \tilde{d})$ becomes a non-uniformly expanding shift with respect to $\mu = \iota_* \mu_\phi$. Moreover, (σ, μ_ϕ) satisfies the requirements of Theorem C. This construction of non-uniformly expanding shifts can also be used to conjugate multidimensional non-uniformly expanding maps with a Markov partition (e.g. [18]*).*

In the next example we show how the argument used in the proof of Theorem D can be used to prove large deviations for hitting times associated to weak Gibbs measures.

Example 5.5. *Let $\sigma : \Sigma_2 \to \Sigma_2$ be the full one-sided shift and μ be a σ-invariant probability measure that is ψ-mixing with respect to the natural partition $\mathcal{P}$ on cylinders: there exists a function $\psi : \mathbb{N} \to \mathbb{R}^+$ so that $\lim_{k\to\infty} \psi(k) = 0$ and satisfying*

$$1 - \psi(k) \leq \frac{\mu(U \cap \sigma^{-n-k}(V))}{\mu(U)\mu(V)} \leq 1 + \psi(k)$$

for every $U \in \mathcal{P}^n$ and $V \in \bigcup_{j=1}^{\infty} \mathcal{P}^j$. These class of measures are a strict subclass of the weak ψ-mixing measures considered by Haydn and Vaienti [8]*. In particular cylinders decay exponentially fast for ψ-mixing measures, i.e., there exists $\eta \in (0,1)$ so that $\mu(\mathcal{P}^{(n)}(x)) \leq \eta^n$ for every $n \geq 1$ (see* [8, Lemma 3]*). Moreover, they considered the so called slow returns $\tau_n(x) = \inf\{k \geq 1 : \exists y \in \mathcal{P}^n(x)$ with $\sigma^k(y) \in \mathcal{P}^n(x)\}$ and proved the following large deviations result for $0 < \delta \leq 1$:*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu(\tau_n(x) \leq [\delta n]) < 0. \tag{5.2}$$

In what follows we use the ideas from Theorem D to provide an alternative proof in the case of μ is a weak Gibbs measure. Notice that $\tau_n(x)$ is a constant on each element of $\mathcal{P}^n$ and $(-\log \mu(\mathcal{P}^{(n)}(x)))_{n\geq 1}$ is a sequence of continuous observables and satisfies that

$$\log(1 - \psi(k)) \leq \log \mu(\mathcal{P}^n \cap \sigma^{-n-k}\mathcal{P}^\ell) - \log \mu(\mathcal{P}^n) - \log \mu(\mathcal{P}^\ell) \leq \log(1 + \psi(k))$$

for every $n, k, \ell \geq 0$, by taking $U = \mathcal{P}^n$ and $V = \mathcal{P}^\ell$ and $\log(1 \pm \psi(\cdot))$ is clearly a bounded function by construction. We use an argument similar to one in the proof of Theorem D. For any $\zeta > 0$ small, if μ is a weak Gibbs measure for the full shift *with respect to a sequence of continuous functions* $\Phi = \{\varphi_n\}$ *then*

$$\mu\Big(\frac{\tau_n}{n} < 1-\varepsilon\Big) = \mu(\tau_n \le [(1-\varepsilon)n]) = \sum_{k=1}^{[(1-\varepsilon)n]} \sum_{\{Q\in\mathcal{P}^n : \tau_n(Q)=k\}} \mu(Q)$$
$$\le K_n e^{-Pn} \sum_{k=1}^{[(1-\varepsilon)n]} \sum_{\{Q\in\mathcal{P}^n : \tau_n(Q)=k\}} e^{\varphi_n(Q)}$$
$$\le e^{(-P+\overline{CP}+\frac{1}{n}\log K_n+\zeta)n}$$

for all large n *(here* $\overline{CP} = \overline{CP}(\sigma, \Phi, \{\tau_n \le [(1-\varepsilon)n]\})$ *and* $P = P_{\mathrm{top}}(\sigma,\Phi)$*). Hence, the convergence is exponential provided* $\overline{CP} < P$*. For instance, if* σ *is the full shift on two symbols and* $\varphi_n \equiv 0$ *for all* n *then*

$$\overline{CP} \le \limsup_{n\to\infty} \frac{1}{n} \log 2^{[[1-\varepsilon)n]} < \log 2,$$

which implies the exponential large deviation estimate (5.2).

Acknowledgments. This work was done during a visit of the first author to Soochow University - China, whose research conditions and hospitality are greatly acknowledged. This work is partially supported by the Priority Academic Program Development of Jiangsu Higher Education Institutions. P. Varandas was partially supported by a postdoctoral fellowship by CNPq-Brazil at University of Porto, Portugal. Y. Zhao is the corresponding author and was partially supported by NSFC (11371271). The authors are grateful to T. Bomfim for useful conversations and to the anonymous referee for a careful reading of the manuscript.

References

[1] L. Barreira. Nonadditive thermodynamic formalism: equilibrium and Gibbs measures. *Disc. Contin. Dyn. Syst.*, 16: 279–305, 2006.

[2] T. Bomfim and P. Varandas. Multifractal analysis for weak Gibbs measures: from large deviations to irregular sets *Ergod. Th. & Dynam. Sys.*, to appear 2015.

[3] T. Bomfim and P. Varandas. Multifractal analysis of the irregular set for almost-additive sequences via large deviations. *Preprint arXiv: 1410.2220*, 2014.

[4] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lect. Notes in Math.* Springer Verlag, 1975.

[5] Y. Cao, D. Feng and W. Huang. The thermodynamic formalism for sub-additive potentials. *Discrete Contin. Dyn. Syst.*, 20: 639–657, 2008.

[6] T. Downarowicz. *Entropy in dynamical systems.* New Mathematical Monographs, 18. Cambridge University Press, Cambridge, 2011

[7] D. Feng, W. Huang, Lyapunov spectrum of asymptotically sub-additive potentials. *Commun. Math. Phys.*, 297: 1–43, 2010.

[8] N. Haydn and S. Vaienti. The Rényi entropy function and the large deviation of short return times. *Ergodic Theory Dynam. Systems*, 30, no. 1, 159–179, 2010.

[9] G. Iommi, Y. Yayama, Weak Gibbs measures as Gibbs measures for aysmptotically additive sequences. Arxiv: 1505.00977

[10] T. Jordan and M. Rams. Multifractal analysis of weak Gibbs measures for non-uniformly expanding C^1 maps. Ergodic Theory Dynam. Systems, 31(1): 143–164, 2011.

[11] A. Katok. Fifty years of Entropy in Dynamics: 1958 - 2007. *Journal of Modern Dynamics*, 1:545–596, 2007.

[12] Y. Katznelson and B. Weiss. A simple proof of some ergodic theorems. *Israel Journal of Math.*, 42:4, 291–296,1982.

[13] V. Kleptsyn, D. Ryzhov and S. Minkov. Special ergodic theorems and dynamical large deviations. *Nonlinearity* 25, 3189-3196, 2012.

[14] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.*, 260, no. 1, 131–146, 2005.
[15] A. Mummert. The thermodynamic formalism for almost-additive sequences. *Discrete Contin. Dyn. Syst.*, 16: 435–454, 2006.
[16] W. Philipp and W. Stout. Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.*, 2, no.161, 1975.
[17] J. Rousseau, P. Varandas and Y. Zhao. Entropy formula for dynamical systems with mistakes. *Discret. Contin. Dyn. Syst.*, 32:12, 4391–4407, 2012.
[18] P. Varandas. Correlation decay and recurrence asymptotics for some robust nonuniformly hyperbolic maps. *J. Stat. Phys.*, 133 (2008), no. 5, 813–839.
[19] P. Varandas, Entropy and Poincaré recurrence from a geometrical viewpoint, *Nonlinearity*, 22: 2365–2375, 2009.
[20] P. Varandas and Y. Zhao. Weak specification properties and large deviations for non-additive potentials. *Ergod Th. & Dynam. Sys.*, 35:3, 968–993, 2015.
[21] P. Walters, An introduction to ergodic theory, *Springer-Verlag*, New York, 1982.
[22] Y. Zhao, L. Zhang and Y. Cao. The asymptotically additive topological pressure on the irregular set for asymptotically additive potentials. *Nonlinear Analysis* 74: 5015C5022, 2011.

Departamento de Matemática, Universidade Federal da Bahia, Av. Ademar de Barros s/n, 40170-110 Salvador, Brazil. & CMUP, University of Porto, Portugal
E-mail address: paulo.varandas@ufba.br
URL: www.pgmat.ufba.br/varandas

Departament of Mathematics, Soochow University, Suzhou 215006, Jiangsu, P.R. China
E-mail address: zhaoyun@suda.edu.cn