Ergod. Th. & Dynam. Sys. (2010), **30**, 159–179
doi:10.1017/S0143385709000030

The Rényi entropy function and the large deviation of short return times

NICOLAI HAYDN† and SANDRO VAIENTI‡§

† *Department of Mathematics, University of Southern California, Los Angeles, CA 90089-1113, USA*
(e-mail: nhaydn@math.usc.edu)
‡ *Centre de Physique Théorique, UMR 6207, CNRS, Luminy Case 907, F-13288 Marseille Cedex 9, France*
§ *Universities of Aix-Marseille I, II and Toulon-Var, Fédération de Recherche des Unités de Mathématiques de Marseille, France*
(e-mail: vaienti@cpt.univ-mrs.fr)

(*Received* 31 *March* 2008 *and accepted in revised form* 9 *December* 2008)

Abstract. We consider the Rényi entropy function for weakly ψ-mixing systems. The first main result of the paper establishes existence and regularity properties. The second main result is obtaining the decay rate for the large deviation of the return time to cylinder sets; we show that it is exponential with a rate given by the Rényi entropy function. Finally, we also obtain bounds for the free energy.

1. *Introduction*

The Rényi entropies [**34**] were extensively studied in the 1980s because of their connections with various generalized spectra for dimensions of dynamically invariant sets; see, for instance, [**9, 10, 16, 18, 22, 23, 31**]. The commonly adopted definition generalizes the usual measure-theoretic entropy. Let T be a transformation on a measurable space Ω, μ a T-invariant probability measure on Ω and $R_{\mathcal{A}}$ the *Rényi entropy function* (defined in (4) below) associated with a finite measurable partition $\mathcal{A}$ of Ω.

Up to now, the existence of the Rényi entropy has been proved only in a few special situations: for Bernoulli measures, Markov measures and, more generally, Gibbs measures with Hölder continuous potentials ϕ. In fact, in these cases the Rényi entropies can be expressed easily in terms of the topological pressure $P(\phi)$ of ϕ (see §2.2 below) independently of the partition $\mathcal{A}$ (provided it is generating). Also, Luczak and Szpankowski [**29**] proved the existence of the Rényi entropy for instantly ϕ-mixing measures (i.e. where $\phi(0) < 1$; see below).

The first main result of our paper, Theorem 1, shows existence of the limit (4) for a large class of measures (namely, *dynamically weakly ψ-mixing measures*; see §2.1). Moreover, we prove that for $t \to 0^+$, the entropy $R_{\mathcal{A}}(t)$ converges to the metric-theoretic entropy $h(\mu)$ and that the function $tR_{\mathcal{A}}(t)$ is locally Lipschitz continuous.

In [**36**], Takens and Verbitsky suggested defining the Rényi entropy of order t of the measure-preserving transformation T as the function $\hat{R}(t) = \sup_{\mathcal{A}} R_{\mathcal{A}}(t)$, where the supremum is taken over all finite partitions $\mathcal{A}$ of Ω. This ensures that $\hat{R}(t)$ is a measure-theoretic invariant but, at the same time, becomes trivial, since it was shown that for ergodic measures μ, the function $\hat{R}(t)$ is (for all $t > 0$) identically equal to the entropy $h(\mu)$. In order to 'extract new information about the dynamics from the generalized entropies' [**36**], Takens and Verbitsky introduced the *correlation entropies* by replacing cylinders with dynamical (Bowen) balls. The main application of correlation entropies has been the complete characterization of the multifractal spectrum of local entropies for expansive homeomorphisms with specification [**37**] (see also [**8, 33**] for another approach). In fact, in the latter case the correlation entropies coincide with the Rényi entropies $R_{\mathcal{A}}(t)$ computed with respect to any generating partition $\mathcal{A}$.

The dependency of the Rényi entropy on the partition reflects some of the mixing properties of the system, as can be seen from Theorem 4, where the behavior of $R_{\mathcal{A}}(t)$ for large values of t is related to the frequency of very short returns which, in turn, expresses the way in which the partition models the periodic behavior of T.

In the remainder of this paper, we denote the Rényi entropy simply by $R(t)$, assuming that a given finite generating partition $\mathcal{A}$ has been chosen once and for all. The second main result of this paper (Theorem 4 together with Corollary 5) uses the Rényi entropies to compute the large deviations of the first returns of cylinders A_n of length n. For this purpose, let us introduce the return-times function

$$\tau_A(x) = \min\{k \geq 1 \mid T^k x \in A\},$$

which is finite for μ-almost every $x \in A$ (Poincaré's theorem) and has expectation (on A) equal to 1 (Kac's theorem) when μ is ergodic.

For $n = 1, 2, \ldots,$ let us put $\tau_n(x) = \min_{y \in A_n(x)} \tau_{A_n(x)}(y)$, where $A_n(x)$ denotes the n-cylinder that contains x. This quantity has arisen in several different contexts.

- Since it controls the short returns, it plays a crucial role in establishing the asymptotic (exponential) distribution of the return-times function $\tau_A(x)$ when the measure of the set A goes to zero [**1–3, 21, 24–26**].
- It has been used to define the *recurrence dimension*, serving as the gauge set function for constructing a suitable Carathéodory measure [**5, 7, 32**].
- It has been related to 'algorithmic information content' in [**12**].

The first result on the asymptotic behavior of $\tau_n(x)$ was proved in [**35**] (see also [**6**] for a different proof): for an ergodic measure μ of positive metric entropy $h(\mu)$, we have

$$\liminf_{n\to\infty} \frac{\tau_n(x)}{n} \geq 1 \tag{1}$$

for μ-almost every $x \in \Omega$. For systems which enjoy the specification property, the above limit exists and is equal to 1 almost everywhere [**6, 35**]. The same result holds for a large class of maps on the interval with indifferent fixed points [**20**].

The situation changes considerably for systems with zero entropy. In general, the limit (1) does not exist anymore, and the values of the lim inf and lim sup depend upon the arithmetic properties of the map; see [**14, 27, 28**] for a careful investigation of Sturmian shifts and substitutive systems.

We will prove in §3 that the limit (1) exists almost surely and is equal to 1 even for weakly ψ-mixing measures. This leads immediately to the natural question of computing the large deviations for the process $\tau_n(x)/n$, that is, to check the existence of the limit defining the lower deviation function

$$\lim_{n\to\infty} \frac{1}{n} \log \mu(x;\, \tau_n(x) \leq [\delta n]) \tag{2}$$

for $\delta \leq 1$. The case $\delta > 1$ is not interesting: it gives the value 0 to the above limit since $\tau_n(x) \leq n + \Delta$, where Δ is a constant independent of x and n; see §3.

The existence of the lower deviation function (2) was first established in [**4**] for classical ψ-mixing measures. These are special cases of the measures considered in this paper: they must satisfy the stronger mixing condition

$$\left| \frac{\mu(U \cap T^{-n-k} V)}{\mu(U)\mu(V)} - 1 \right| \leq \psi(k) \tag{3}$$

for all U in $\sigma(\mathcal{A}^n)$, all n and all $V \in \sigma(\mathcal{A}^*)$ (the σ-algebra generated by $\mathcal{A}^n$), where $\mathcal{A}^* = \bigcup_{j=1}^{\infty} \mathcal{A}^j$ (compare this with the definition of weakly ψ-mixing measures given in §2 below). The rate function $\psi(k)$, $k \geq 0$, must converge to zero; moreover, to achieve the existence of the limit (2), the additional assumption $\psi(0) < 1$ was required in [**4**]. This, in particular, implies (see [**4**, Lemma 2.1]) that after having coded the elements of the initial finite partition (of cardinality $|\mathcal{A}| = M$, say) $\mathcal{A} = \bigcup_{i=1}^{M} A_i$ over the alphabet $G = \{1, 2, \ldots, M\}$, for every string $\{i_0, \ldots, i_{n-1}\} \in G^n$ ($n \geq 1$) the cylinder $A_{i_0} \cap T^{-1} A_{i_1} \cap \cdots \cap T^{(n-1)} A_{i_{n-1}}$ has positive measure, which essentially means that the grammar associated to the coding is complete. We shall have no need of this condition, not even for our larger class of weakly ψ-mixing measures. The key result in [**4**] was to relate the lower deviation function to the Rényi entropies for any ψ-mixing measure satisfying the condition $\psi(0) < 1$, but in that paper the existence of the Rényi entropies was assumed, since no general result was known then.

It is well-known that the deviation function can be computed as the Legendre transform of the free energy of the process, provided that the free energy exists and is differentiable with respect to the parameter β (see (5) below). We show in §4 that this is not the case for our process: the free energy will be continuous but not differentiable at the point $\beta = -\gamma_\mu$, where γ_μ is the exponential decay rate of the measures of n-cylinders from Theorem 1. Even if the free energy is not differentiable, one can still derive an upper bound for the lower deviation function, which we shall show to be consistent with the rigorous expression for the lower deviation function in terms of the Rényi entropies. It is interesting to note that the free energy function was also computed in [**4**], but the proof there needed an additional assumption, namely, the existence of a sequence of cylinders with measures decaying exponentially to zero at a rate which is exactly the constant γ_μ, and whose first return is sublinear. We do not need this hypothesis, since we will prove the existence of such a sequence in full generality.

2. *Rényi entropy function*

2.1. *Existence and regularity.* Let T be a transformation on the measurable space Ω and μ a T-invariant probability measure on Ω. Assume that Ω has a finite measurable partition $\mathcal{A}$ whose joins we denote by $\mathcal{A}^k = \bigvee_{j=0}^{k-1} T^{-j}\mathcal{A}$, $k = 1, 2, \ldots$ (the elements of $\mathcal{A}^k$ are commonly referred to as n-cylinders). We assume that $\mathcal{A}$ is generating, i.e. that the elements of $\mathcal{A}^\infty$ are single points. For $t > 0$, we put

$$Z_n(t) = \sum_{A_n \in \mathcal{A}^n} \mu(A_n)^{1+t}$$

and define the *Rényi entropy function* $R_{\mathcal{A}}$ with respect to the partition $\mathcal{A}$ by

$$R_{\mathcal{A}}(t) = \liminf_{n\to\infty} \frac{1}{tn} |\log Z_n(t)|. \tag{4}$$

This section is devoted to proving existence of the limit under some weak mixing conditions. Since the partition $\mathcal{A}$ is given, we will simply write $R(t)$ for the Rényi entropy. We say that the T-invariant probability measure μ on Ω is *weakly ψ-mixing* with respect to the (finite) partition $\mathcal{A}$ if there exist positive functions $\psi^-, \psi^+ : \mathbb{N} \to \mathbb{R}^+$, with $\psi^-(k) < 1$ for all $k \geq \Delta_0$ for some Δ_0, such that

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

for all U in $\sigma(\mathcal{A}^n)$, all n and all $V \in \sigma(\mathcal{A}^*)$ (where $\mathcal{A}^* = \bigcup_{j=1}^\infty \mathcal{A}^j$). From now on, we assume that the measure μ on Ω is a T-invariant non-atomic probability measure which is weakly ψ-mixing, with the functions $1 - \psi^-$ and $1 + \psi^+$ being subexponential, i.e. $\limsup_{k\to\infty}(1/k)|\log(1 - \psi^-(k))| = 0$ and $\limsup_{k\to\infty}(1/k)\log(1 + \psi^+(k)) = 0$. Lemma 3 shows that the measures of cylinder sets decay exponentially fast. Classical ψ-mixing measures correspond to the special case where $\psi^-(k) = \psi^+(k) = \psi(k)$ and $\psi(k) \searrow 0$ as $k \to \infty$ [**11, 15, 17**]. The classical ψ-mixing property implies, in particular, that μ cannot have any atoms.

Put $b_n = \max_{A_n \in \mathcal{A}^n} \mu(A_n)$ and let $\gamma_\mu = \liminf_n (1/n)|\log b_n|$ be *the exponential decay rate of the measures of n-cylinders*. We will now establish the following properties of the Rényi entropy.

THEOREM 1. *Assume that the (non-atomic) measure μ is weakly ψ-mixing and that the functions $1 - \psi^-$ and $1 + \psi^+$ are subexponential. Then:*

(I) *the limit $R(t) = \lim_{n\to\infty}(1/(tn))|\log Z_n(t)|$ exists for $t > 0$, with convergence being uniform for t on compact subsets of $\mathbb{R}^+$;*

(II) *the function $W(t) = tR(t)$ is locally Lipschitz continuous;*

(III) *$R(0) = \lim_{t\to 0^+} R(t) = h(\mu)$;*

(IV) *$R(t)$ is monotonically decreasing on $(0, \infty)$ and $R(t) \to \gamma_\mu$ as $t \to \infty$, where $\gamma_\mu = \liminf_{n\to\infty}(1/n)|\log b_n|$ is positive.*

2.2. *Examples.* *(I) Bernoulli shift.* If Ω is the full shift space over a finite alphabet $\{1, 2, \ldots, M\}$, σ is the left shift transformation, the partition $\mathcal{A}$ is the collection of one-element cylinders and the invariant probability measure μ is given by a probability vector $\vec{p} = (p_1, p_2, \ldots, p_M)$ (where $\sum_i p_i = 1$, $p_i > 0$), then $Z_n(t) = (\sum_i p_i^{1+t})^n$ and the Rényi entropy is $R(t) = (1/t)\log \sum_i p_i^{1+t}$ for $t > 0$ and equals the metric entropy $h_\mu = \sum_i p_i |\log p_i|$ for $t = 0$.

(II) Markov chains. Again Ω is the shift space over the alphabet $\{1, 2, \ldots, M\}$ and $\mathcal{A}$ is the usual partition of one-element cylinders. The invariant probability measure μ is now given by an $M \times M$ stochastic matrix P (we assume P is irreducible) and probability vector $\vec{p}$, with $\vec{p}P = \vec{p}$ and $P1 = 1$. The cylinder set $U(x_1 \ldots x_n) \in \mathcal{A}^n$ given by the n-word $x_1 x_2 \ldots x_n$ then has the measure $\mu(x_1 \ldots x_n) = p_{x_1} P_{x_1 x_2} P_{x_2 x_3} \cdots P_{x_{n-1} x_n}$. Hence

$$Z_n(t) = \sum_{x_1 x_2 \ldots x_n} p_{x_1}^{1+t} P_{x_1 x_2}^{1+t} \cdots P_{x_{n-1} x_n}^{1+t},$$

where the sum is over all (admissible) n-words. The non-negative $M \times M$-matrix $P(t)$ whose entries are $P_{ij}(t) = P_{ij}^{1+t}$ has, by the Perron–Frobenius theorem, a single largest positive eigenvalue λ_t and a strictly positive (and normalized) left eigenvector $\vec{w}(t)$. (Note that λ_t is a continuous function of t and $\lambda_0 = 1$.) Thus (with $p_i(t) = p_i^{1+t}$, $i = 1, \ldots, M$),

$$\lambda_t^{-n} \vec{p}(t) P(t)^n \to (\vec{p} \cdot \vec{w}(t)) \vec{w}(t)$$

(exponentially fast) as $n \to \infty$. We therefore obtain that $R(t) = (1/t)\log \lambda_t$ if t is positive and that $R(0) = h_\mu = \sum_{ij} p_i P_{ij} |\log P_{ij}|$ if $t = 0$.

(III) Gibbs measures. [**10, 36**] If μ is a Gibbs measure for the potential function ϕ [**13**], then the Rényi entropy $R(t) = (1/t)\big((1+t)P(\phi) - P((1+t)\phi)\big)$ (where P is the pressure function) is analytic for $t > 0$.

2.3. *Proof of Theorem 1.* To prove Theorem 1, we need the following technical lemma about ψ-mixing measures. The notation $\psi^\pm$ means that ψ^+ applies when the argument inside the absolute value is positive and ψ^- applies when the argument inside the absolute value is negative.

LEMMA 2. *Assume there are sets $B_j \in \sigma(\mathcal{A}^{n_j})$, $j = 1, 2, \ldots, k$, for some integers n_j. If μ is weakly ψ-mixing, then*

$$\left| \mu\bigg(\bigcap_{j=1}^{k} T^{-N_j} B_j \bigg) - \prod_{j=1}^{k} \mu(B_j) \right| \le \Big(\big(1 + \psi^\pm(\Delta)\big)^{k-1} - 1 \Big) \prod_{j=1}^{k} \mu(B_j),$$

for all $\Delta \ge 0$, where $N_j = n_1 + n_2 + \cdots + n_{j-1} + (j-1)\Delta$ with $N_0 = 0$.

Proof. For $\ell = 1, 2, \ldots, k$, put

$$D_\ell = \bigcap_{j=\ell}^{k} T^{-(N_j - N_\ell)} B_j.$$

In particular, note that $\bigcap_{j=1}^{k} T^{-N_j} B_j = D_1$, $D_k = B_k$ and

$$D_\ell = B_\ell \cap T^{-n_\ell - \Delta} D_{\ell+1}.$$

By the mixing property, $|\mu(D_\ell) - \mu(B_\ell)\mu(D_{\ell+1})| \le \psi^{\pm}(\Delta)\mu(B_\ell)\mu(D_{\ell+1})$; upon repeated application this yields, by the triangle inequality,

$$\left|\mu\left(\bigcap_{j=1}^{k} T^{-N_j}B_j\right) - \prod_{j=1}^{k}\mu(B_j)\right| \le \psi^{\pm}(\Delta)\sum_{\ell=1}^{k-1}\mu\left(\bigcap_{j=1}^{\ell-1} T^{-N_j}B_j\right)\prod_{j=\ell}^{k-1}\mu(B_j)$$
$$\le ((1+\psi^{\pm}(\Delta))^{k-1}-1)\prod_{j=1}^{k}\mu(B_j). \qquad \square$$

LEMMA 3. *There exists a constant $\eta \in (0, 1)$ such that $\mu(A_n) \le \eta^n$ for all $A_n \in \mathcal{A}^n$ and all n.*

Proof. Fix a $\Delta > 0$ and an m so that $b_m = \max_{A_m \in \mathcal{A}^m}\mu(A_m) \le (1/2)(1+\psi^+(\Delta))^{-1}$ (note that $b_m \searrow 0$ as $m \to \infty$, since μ has no atoms). Then, for any n (large) and $A_n \in \mathcal{A}^n$, one has $A_n \subset \bigcap_{j=0}^{k-1} T^{-km'}A_m(T^{jm'}A_n)$, where $k = [n/m']$, $m' = m + \Delta$ and $A_m(T^{jm'}A_n)$ is the m-cylinder that contains $T^{jm'}A_n$ $(j \le k-1)$. By Lemma 2,

$$\mu(A_n) \le \mu\left(\bigcap_{j=0}^{k-1} T^{-km'}A_m(T^{jm'}A_n)\right) \le (1+\psi^+(\Delta))^{k-1}\prod_{j=1}^{k}\mu(A_m(T^{jm'}A_n))$$
$$\le (1+\psi^+(\Delta))^k\, b_m^k \le 2^{-k}.$$

Hence $\eta \le 2^{-1/m'}$. $\square$

Remark 1. The exponential decay of cylinders implies, in particular, that the metric entropy of a weakly ψ-mixing measure μ is positive. In fact, $h(\mu) \ge |\log\eta| > 0$.

Proof of (I). To prove that the limit that defines $R(t)$ exists, we will show that the sequence $a_n = |\log Z_n(t)|$ is 'nearly' subadditive; a standard argument then ensures that the limit exists.

Let m and $\Delta \ge \Delta_0$ (the 'gap') be integers, put $m' = m + \Delta$ and let $n = km' - \Delta$ be a large integer. Put $\tilde{\mathcal{A}}^n = \bigvee_{j=0}^{k-1} T^{-jm'}\mathcal{A}^m$ and define, for some $\beta > 1$,

$$\mathcal{G}_n = \{A_n \in \mathcal{A}^n \mid \mu(A_n) \ge e^{-k\Delta^\beta}\mu(\tilde{A}_n)\},$$

where $\tilde{A}_n = \bigcap_{j=0}^{k-1} T^{-km'}A_m(T^{jm'}A_n)$. Then, for every A_n, one has

$$\mu\left(\bigcup_{A'_n \subset \tilde{A}_n, A'_n \in \mathcal{G}_n} A'_n\right) = \mu(\tilde{A}_n) - \mu\left(\bigcup_{A'_n \subset \tilde{A}_n, A'_n \notin \mathcal{G}_n} A'_n\right) \ge (1 - |\mathcal{A}|^{k\Delta}e^{-k\Delta^\beta})\mu(\tilde{A}_n),$$

as $\tilde{A}_n = \bigcup_{A'_n \subset \tilde{A}_n, A'_n \in \mathcal{A}^n} A'_n$ has k 'gaps', each of which is of length Δ. This implies that if $|\mathcal{A}|^{\Delta}e^{-\Delta^\beta} < 1$, then for every $\tilde{A}_n \in \tilde{\mathcal{A}}^n$ there exists an $A'_n \subset \tilde{A}_n$, $A'_n \in \mathcal{A}^n$, which also belongs to $\mathcal{G}_n$. As $\Delta \ge \Delta_0$, we get

$$Z_n(t) = \sum_{A_n \in \mathcal{A}^n}\mu(A_n)^{1+t} \ge e^{-k\Delta^\beta(1+t)}\sum_{\tilde{A}_n \in \tilde{\mathcal{A}}^n}\mu(\tilde{A}_n)^{1+t}$$
$$= e^{-k\Delta^\beta(1+t)}Z_m(t)^k((1+\mathcal{O}(\psi^-(\Delta)))^{k-1})^{1+t},$$

where we have used the fact that $\mu(\tilde{A}_n) = (1+\mathcal{O}(\psi^{\pm}(\Delta)))^{k-1}\prod_{j=0}^{k-1}\mu(A_m(T^{jm'}\tilde{A}_n))$ (the mixing property). Hence we obtain

$$\begin{aligned}|\log Z_n(t)| &\le k|\log Z_m(t)| + k\Delta^{\beta}(1+t) + (1+t)|\log(1-\psi^{-}(\Delta))^{k-1}|\\ &\le k|\log Z_m(t)| + \mathcal{O}(k\Delta^{\beta}(1+t)).\end{aligned}$$

If we put $a_n = |\log Z_n(t)|$, then $a_n \le ka_m + ck\Delta^{\beta}$ and

$$\frac{a_{km'}}{km'} \le \frac{a_m}{m'} + c\frac{\Delta^{\beta}}{m'} = \frac{m}{m'}\frac{a_m}{m} + c\frac{\Delta^{\beta}}{m'}.$$

If we put $\Delta \sim m^{\alpha}$ so that $\alpha\beta < 1$, then

$$\limsup_n \frac{a_n}{n} \le \frac{m}{m+\Delta}\frac{a_m}{m} + \mathcal{O}\bigg(\frac{\Delta^{\beta}}{m}\bigg) \quad \text{for all } m.$$

Hence $\limsup_n(a_n/n) \le \liminf_m(a_m/m)$.

We also have

$$Z_n(t) \le |\mathcal{A}|^{k\Delta} Z_m(t)^k(1+\mathcal{O}(\psi^{+}(\Delta)))^{(k-1)(1+t)},$$

which implies that

$$|\log Z_n(t)| \ge k|\log Z_m(t)| + \mathcal{O}(k\Delta(1+t)).$$

This ensures uniform convergence for t in compact subsets of $\mathbb{R}^{+}$.

Proof of (II). For $t > 0$, let us put $H_n(t) = \sum_{A_n\in\mathcal{A}^n}\mu(A_n)^{1+t}|\log\mu(A_n)|$; clearly, $h(\mu) = \lim_{n\to\infty} H_n(0)/n$ and $dZ_n(t)/dt = H_n(t)$. We will show that the sequence H_n has an additive-like behavior in the sense that $H_{km} = \mathcal{O}(km)$, where the implied constant is bounded and bounded away from 0, uniformly in k and m. This then implies the Lipschitz property.

As above, let $\tilde{\mathcal{A}}^n = \bigvee_{j=0}^{k-1} T^{-jm'}\mathcal{A}^m$ and, in order to cut k gaps of length Δ, put

$$\mathcal{G}_n = \{A_n \in \mathcal{A}^n \mid \mu(A_n) \ge e^{-k\Delta^{\beta}}\mu(\tilde{A}_n)\}$$

for some $\beta > 1$, where $\tilde{A}_n \in \tilde{\mathcal{A}}^n$ is such that $A_n \subset \tilde{A}_n$ and $n = km' - \Delta$ $(m' = m+\Delta)$. The sum over $\mathcal{A}^n$ that defines Z_n is split into two parts: (i) over $\mathcal{G}_n$ and (ii) over the complement of $\mathcal{G}_n$.

(i) On the set $\mathcal{A}^n \setminus \mathcal{G}_n$ we have $\mu(A_n) \le e^{-k\Delta^{\beta}}\mu(\tilde{A}_n)$, where $A_n \in \mathcal{G}_n$ with $A_n \subset \tilde{A}_n \in \tilde{\mathcal{A}}^n$. Choose $\gamma \in (1,\beta)$ and let $\mathcal{G}'_n = \{A'_n \in \mathcal{A}^n \mid \mu(A'_n) \ge e^{-k\Delta^{\gamma}}\mu(\tilde{A}_n)t\}$. Then we get that for all $A_n \notin \mathcal{G}_n$,

$$\mu(A_n) \le e^{-k\Delta^{\beta}}\mu(\tilde{A}_n) \le e^{-k\Delta^{\beta}}e^{-k\Delta^{\gamma}}\mu(A'_n),$$

where $A'_n \in \mathcal{G}'_n$ is such that $A'_n \subset \tilde{A}_n$ (such an A'_n exists since $|\mathcal{A}|^{\Delta}e^{-\Delta^{\gamma}} < 1$ for Δ large enough). Thus

$$\begin{aligned}\sum_{A_n\notin\mathcal{G}_n}|\log\mu(A_n)|\mu(A_n)^{1+t} &\le e^{-k(1+t)\Delta^{\beta}}\sum_{A_n\notin\mathcal{G}_n}|\log\mu(A_n)|\mu(\tilde{A}_n)^{1+t}\\ &\le e^{-k(1+t)(\Delta^{\beta}-\Delta^{\gamma})}\sum_{A'_n\in\mathcal{G}'_n}|\log\mu(A'_n)|\mu(A'_n)^{1+t}\\ &\le e^{-k(1+t)(\Delta^{\beta}-\Delta^{\gamma})}H_n.\end{aligned}$$

(ii) If $A_n \in \mathcal{G}_n$, then $\log \mu(A_n) = \log \mu(\tilde{A}_n) + \mathcal{O}(k\Delta^\beta)$, and we obtain

$$\begin{aligned}
H_n(t) &= \sum_{A_n \in \mathcal{A}^n} |\log \mu(A_n)|\mu(A_n)^{1+t} \\
&= \sum_{A_n \in \mathcal{G}_n} (|\log \mu(\tilde{A}_n)| + \mathcal{O}(k\Delta^\beta))\mu(A_n)^{1+t} + \sum_{A_n \notin \mathcal{G}_n} |\log \mu(A_n)|\, \mu(A_n)^{1+t} \\
&= \sum_{A_n \in \mathcal{G}_n} |\log \mu(\tilde{A}_n)|\mu(A_n)^{1+t} + \mathcal{O}(k\Delta^\beta) Z_n + \mathcal{O}(e^{-k(1+t)(\Delta^\beta - \Delta^\gamma)}) H_n,
\end{aligned}$$

where in the last step we have used the estimate from part (i).

The mixing property $\mu(\tilde{A}_n) = (1 + \mathcal{O}(\psi^\pm(\Delta)))^{k-1} \prod_{j=0}^{k-1} \mu(A_m \circ T^{jm'})$ is applied to the principal term, giving

$$\sum_{A_n \in \mathcal{A}^n} |\log \mu(\tilde{A}_n)|\mu(A_n)^{1+t} = \sum_{j=0}^{k-1} X^j + \mathcal{O}(k(\psi^-(\Delta) + \psi^+(\Delta))),$$

where $X^j = \sum_{A_n \in \mathcal{A}^n} |\log \mu(A_m \circ T^{jm'})|\mu(A_n)^{1+t}$. To further examine X^j, let us put

$$\tilde{\mathcal{A}}_j^n = \mathcal{A}^{jm'-\Delta} \vee T^{-jm'}\mathcal{A}^m \vee T^{-(j+1)m'-\Delta}\mathcal{A}^{n-(j+1)m'-\Delta},$$

where we have opened up two gaps of length Δ (with $\Delta \geq \Delta_0$), the first after j blocks and the second after $j+1$ blocks ($j = 0, \dots, k-1$), with the obvious modification that for $j = 0$ or $k-1$ there is only a single gap. We now put

$$\mathcal{G}_n^j = \{A_n \in \mathcal{A}^n \mid \mu(A_n) \geq e^{-\Delta^\beta}\mu(\tilde{A}_n^j)\},$$

where $\tilde{A}_n^j \in \tilde{\mathcal{A}}_j^n$ is such that $A_n \subset \tilde{A}_n^j$. The sum in X^j over $\mathcal{A}^n$ is split into two parts: (a) over $\mathcal{G}_n^j$ and (b) over its complement $\mathcal{A}^n \setminus \mathcal{G}_n^j$.

(a) For the sum over $\mathcal{G}_n^j$, the mixing property

$$\mu(\tilde{A}_n^j) = (1 + \mathcal{O}(\psi(\Delta)))\mu(A_{jm'-\Delta})\mu(A_m \circ T^{jm'})\mu(A_{n-(j+1)m'-\Delta} \circ T^{-(j+1)m'-\Delta})$$

for $\tilde{A}_n^j \in \mathcal{A}_j^n$ yields

$$\begin{aligned}
&\sum_{A_n \in \mathcal{G}_n^j} |\log \mu(A_m \circ T^{jm'})|\mu(A_n)^{1+t} \\
&\quad \in [e^{-(1+t)\Delta^\beta}, |\mathcal{A}|^{2\Delta}] \sum_{\tilde{A}_n \in \tilde{\mathcal{A}}_j^n} |\log \mu(A_m \circ T^{jm'})|\mu(\tilde{A}_n^j)^{1+t} \\
&\quad = [e^{-(1+t)\Delta^\beta}, |\mathcal{A}|^{2\Delta}](1 + \mathcal{O}(\psi^\pm(\Delta)))Z_{jm'-\Delta} H_m Z_{n-(j+1)m'-\Delta}.
\end{aligned}$$

(b) For the sum over $\mathcal{A}^n \setminus \mathcal{G}_n^j$, we estimate as follows:

$$\begin{aligned}
&\sum_{A_n \notin \mathcal{G}_n^j} |\log \mu(A_m \circ T^{jm'})|\mu(A_n)^{1+t} \\
&\quad \leq |\mathcal{A}|^{2\Delta} e^{-(1+t)\Delta^\beta} \sum_{\tilde{A}_n^j \in \mathcal{A}_j^n} |\log \mu(A_m \circ T^{jm'})|\mu(\tilde{A}_n^j)^{1+t} \\
&\quad \leq |\mathcal{A}|^{2\Delta} e^{-(1+t)\Delta^\beta}(1 + \mathcal{O}(\psi^\pm(\Delta)))^{1+t} Z_{jm'-\Delta} H_m Z_{n-(j+1)m'-\Delta}.
\end{aligned}$$

Similarly, one shows that $Z_n \in [|\mathcal{A}|^{2\Delta} e^{-(1+t)\Delta^\beta}, |\mathcal{A}|^{2\Delta}] Z_{jm'-\Delta} Z_m Z_{n-(j+1)m'-\Delta}$. Hence we get

$$H_n \in \left[\frac{1}{c_1}, c_1\right] \sum_{j=0}^{k-1} \frac{Z_{jm'-\Delta} H_m Z_{n-(j+1)m'-\Delta}}{Z_n} + \mathcal{O}(k\Delta^\beta)$$

for some constant $c_1 \approx 2|\mathcal{A}|^{2\Delta} e^{(1+t)\Delta^\gamma}$ and, consequently,

$$H_n \in \left[\frac{1}{c_1^2}, c_1^2\right] k H_m + \mathcal{O}(k\Delta^\beta).$$

This implies that

$$\limsup_{n\to\infty} \frac{1}{n} H_n \le c_1^2 \frac{1}{m} H_m$$

and, similarly,

$$\liminf_{n\to\infty} \frac{1}{n} H_n \ge c_1^{-2} \frac{1}{m} H_m.$$

It follows that $c_2 c_1^{-2}|s| \le W(t+s) - W(t) \le c_2 c_1^2 |s|$ for small s (e.g. $-t \le s \le 1$), for some positive constant c_2 (which is equal to $H_m(t)/m$ for some m).

Proof of (III). To prove existence of the limit as $t \to 0$, we must obtain more delicate estimates for the near-additivity of the sequence $H_n(t)$ for t close to 0.

With $H_n(t) = \sum_{A_n \in \mathcal{A}^n} \mu(A_n)^{1+t} |\log \mu(A_n)|$, as above, we get

$$\begin{aligned}
H_{n+m}(t) &= \sum_{A_{n+m} \in \mathcal{A}^{n+m}} \mu(A_{n+m})^{1+t} \left|\log \frac{\mu(A_{n+m})}{\mu(A_m)} + \log \mu(A_m)\right| \\
&= \sum_{A_{n+m} \in \mathcal{A}^{n+m}} \mu(A_{n+m})^{1+t} |\log \mu(A_m)| \\
&\quad + \frac{1}{1+t} \sum_{A_{n+m} \in \mathcal{A}^{n+m}} \mu(A_{n+m})^{1+t} \left|\log \left(\frac{\mu(A_{n+m})}{\mu(A_m)}\right)^{1+t}\right| \\
&\le \sum_{A_m \in \mathcal{A}^m} \mu(A_m)^{1+t} |\log \mu(A_m)| \\
&\quad + \frac{1}{1+t} \sum_{A_n \in \mathcal{A}^n} Z_m(t) \sum_{A_m \in \mathcal{A}^m} \frac{\mu(A_m)^{1+t}}{Z_m(t)} \phi\left(\left(\frac{\mu(A_{n+m})}{\mu(A_m)}\right)^{1+t}\right),
\end{aligned}$$

where A_{n+m} stands for $A_m \cap T^{-m} A_n$ and $\phi(s) = -s \log s$ is concave on $(0, 1)$ and increasing on $(0, 1/e)$. Thus,

$$\begin{aligned}
H_{n+m}(t) &\le H_m(t) + \frac{Z_m(t)}{1+t} \sum_{A_n \in \mathcal{A}^n} \phi\left(\sum_{A_m \in \mathcal{A}^m} \frac{\mu(A_{n+m})^{1+t}}{Z_m(t)}\right) \\
&\le H_m(t) + \frac{Z_m(t)}{1+t} \sum_{A_n \in \mathcal{A}^n} \phi\left(\frac{\mu(A_n)^{1+t}}{Z_m(t)}\right),
\end{aligned}$$

provided that $\mu(A_n)^{1+t}/Z_m(t) \le 1/e$ for every $A_n \in \mathcal{A}^n$. Hence

$$H_{n+m}(t) \le H_m(t) + \frac{1}{1+t}\sum_{A_n\in\mathcal{A}^n} \mu(A_n)^{1+t}\left|\log\frac{\mu(A_n)^{1+t}}{Z_m(t)}\right|$$
$$= H_m(t) + H_n(t) + \frac{1}{1+t} Z_n(t)|\log Z_m(t)|,$$

as $Z_m \le 1$. Now we apply this estimate repeatedly. In order to satisfy the condition $\mu(A_{jm})^{1+t}/Z_m(t) \le 1/e$ for every $A_{jm} \in \mathcal{A}^{jm}$, $j = 1, \ldots, k$, let us note that the measure of the cylinder sets goes to zero by Lemma 3. Hence, for a given m, we can find an integer J so that $\mu(A_{jm})^{1+t}/Z_m(t) \le 1/e$ for every $A_{jm} \in \mathcal{A}^{jm}$ and all $j > J$. Moreover, since $W(0) = 0$ and $(\log Z_n(t))/n$ converge uniformly to $W(t)$ for $t \in (0, \delta)$ (where $\delta > 0$), we can let $\varepsilon > 0$ and choose $\delta > 0$ so that $|W(t)| < \varepsilon/2$ and N so that $|(\log Z_n(t))/n - W(t)| < \varepsilon/2$, for all $n \ge N$ and $t \in (0, \delta)$. Hence $1 \ge Z_n(t) \ge e^{-\varepsilon n}$ for $n \ge N$ and $t \in (0, \delta)$. Assume that $m > N$. Then we get almost-subadditivity for the sequence $H_n(t)$:

$$H_{km}(t) = H_{Jm}(t) + (k - J)H_m(t) + \mathcal{O}(k\varepsilon m)$$

and consequently, as $k \to \infty$,

$$\lim_{n\to\infty}\frac{H_n(t)}{n} = \frac{H_m(t)}{m} + \mathcal{O}(\varepsilon)$$

for every $m > N$. Therefore, if $t \in (0, \delta)$, then

$$W(t) = \lim_{n\to\infty}\frac{\log Z_n(t)}{n} = \lim_{n\to\infty}\frac{1}{n}\int_0^t H_n(s)\,ds = \frac{1}{m}\int_0^t H_m(s)\,ds + \mathcal{O}(\varepsilon t)$$

and, consequently,

$$R(0) = W'(0) = \lim_{t\to 0^+}\frac{1}{tm}\int_0^t H_m(s)\,ds + \mathcal{O}(\varepsilon) = \frac{1}{m}H_m(0)\,ds + \mathcal{O}(\varepsilon).$$

Since $\varepsilon > 0$ was arbitrary, we get that $R(0) = \lim_{m\to\infty}(H_m(0)/m)$ (we need $m > N_\varepsilon$, where $N_\varepsilon \to \infty$ as $\varepsilon \to 0^+$).

Proof of (IV). The fact that R is decreasing was noted in, for example, [**9, 36**]. Since

$$b_n^{1+t} \le Z_n(t) \le \sum_{A_n\in\mathcal{A}^n}\mu(A_n)b_n^t = b_n^t,$$

we obtain that $|\log b_n|/n \le R(t) \le ((1+t)/t)|\log b_n|/n$ for all n (this estimate is true universally, independently of mixing properties). Hence $\gamma_\mu \le R(t) \le ((1+t)/t)\gamma_\mu$ for all $t > 0$, where γ_μ is strictly positive since, by Lemma 3, $\gamma_\mu \ge |\log\eta| > 0$. □

As Lemma 3 shows, the measure of cylinder sets always decays exponentially fast for weakly ψ-mixing measures. Clearly, if the measure of cylinder sets decays subexponentially (i.e. $\gamma_\mu = 0$), then the Rényi entropy $R(t)$ is identically zero on $(0, \infty)$.

3. *Short return times*

In the introduction, we recalled that for every ergodic measure μ with positive entropy, $\liminf_{n\to\infty}(\tau_n(x)/n) \geq 1$ almost everywhere. Since a weakly ψ-mixing measure μ has positive entropy (see the remark following Lemma 3), we obtain $\liminf_{n\to\infty}(\tau_n(x)/n) \geq 1$ for μ-almost every $x \in \Omega$. In order to get the upper bound, let $x \in \Omega$ and note that by the weak ψ-mixing property,

$$\frac{\mu(A_n(x) \cap T^{-n-\Delta} A_n(x))}{\mu(A_n(x))^2} \geq 1 - \psi^-(\Delta) > 0$$

for $\Delta \geq \Delta_0$. This implies that $\tau_n(x) \leq n + \Delta$ and, since $\Delta \geq \Delta_0$ is fixed, we obtain that $\limsup_{n\to\infty}(\tau_n(x)/n) \leq 1$ for every $x \in \Omega$. Hence

$$\lim_{n\to\infty} \frac{1}{n}\tau_n(x) = 1$$

almost everywhere. This section concerns the large deviations of the process τ_n; in other words, we are interested in the asymptotic behavior of the distributions

$$\mathbb{P}(\tau_n \leq [\delta n]) = \mu(\{x \mid \tau_n(x) \leq [\delta n]\}).$$

Since $\tau_n(x)$ is obviously constant for all points in the same cylinder around x, we can replace the set $\{x \mid \tau_n(x) \leq [\delta n]\}$ by

$$\mathcal{C}_n(\delta) = \{A_n \in \mathcal{A}^n \mid \tau_n(A_n) \leq [\delta n]\},$$

which measures the probability of points having very short returns, where $\tau_n(A_n) = \min\{k \geq 1 \mid A_n \cap T^{-k} A_n \neq \emptyset\} = \tau_n(x)$ for $x \in A_n$. In order to analyze the size of the set, we follow [**4**] and define the sets

$$B_n(j) = \left\{ A_n \in \mathcal{A}^n \;\middle|\; \frac{j}{\tau_n(A_n)} \in \mathbb{N} \right\},$$

where $n \in \mathbb{N}$ and $j = 1, \ldots, n$. Clearly, $B_n(j) \in \sigma(\mathcal{A}^n)$ for all j and, looking at the symbolic representation of the n-cylinders in $B_n(j)$, we note that there are two cases:

(i) if $j \leq n/2$ and x is a point in $B_n(j)$, then the first n symbols of points in it are

$$(x_1 x_2 \ldots x_j)^{n'} x_1 x_2 \ldots x_r,$$

where $n' = [n/j]$ and $r = n - j[n/j]$ $(r < j)$;

(ii) if $j > n/2$ and A_n is an n-cylinder in $B_n(j)$, then the first n symbols of points in it are

$$x_1 x_2 \ldots x_{n-j} x_1 x_2 \ldots x_{2j-n} x_1 x_2 \ldots x_{n-j},$$

where the (remainder) middle portion is of length $n - 2(n - j) = 2j - n$.

Let us put

$$\mathcal{S}_n(\lambda) = \{A_n \in \mathcal{A}^n \mid \tau_n(A_n) = [n\lambda]\}.$$

The purpose of this section is to determine the decay rate of the measure of the set $\mathcal{S}_n(\lambda)$ as n goes to infinity. As λ varies over the unit interval, we obtain the short recurrence spectrum for the measure μ. Let us note that for every n we have that $\mathcal{C}_n(\delta) = \bigcap_{\lambda<\delta} \mathcal{S}_n(\lambda)$.

Recall that $W(t) = tR(t) = \lim_{n\to\infty}(\log Z_n(t))/n$. We define the function $M(\lambda)$ on the interval (0, 1] as follows:

$$M(\lambda) = (1 - \lambda\ell)(W(\ell) - W(\ell - 1)) + \delta W(\ell - 1),$$

where $\ell = [1/\lambda]$; note that $1 - \lambda\ell$ linearly interpolates between the values $1/(k + 1)$ and 0 on the interval $(1/(k + 1), 1/k)$. The function $M(\lambda)$ is continuous on (0, 1) and piecewise affine on the intervals $(1/(1 + k), 1/k)$, and it takes the values $M(1/k) = (1/k)W(k - 1)$ for $k = 1, 2, \ldots$ (in particular, $M(1) = 0$). The function $M(\lambda)$ interpolates the function $\hat{M}(\lambda) = (\lambda/(1 + \lambda))W(1/\lambda)$ between the points $\lambda = 1/k$ for $k = 1, 2, \ldots$. Changing coordinates to $t = (1 + \lambda)/\lambda$, we get $\hat{M}(\lambda) = (1/t)W(t - 1)$. This function is increasing for $t > 1$, as can be seen from the derivatives of the approximating functions. To wit,

$$\frac{d}{dt}\frac{1}{tn}|\log Z_n(t-1)| = \frac{1}{t^2 Z_n(t-1)} \sum_{A_n\in\mathcal{A}^n} \mu(A_n)^t \left|\log \frac{\mu(A_n)^t}{Z_n(t-1)}\right|,$$

which is positive for every n. Since $\lim_{n\to\infty}(1/(tn))|\log Z_n(t - 1)| = (1/t)W(t - 1)$, we conclude that $(1/t)W(t - 1)$ is increasing on $(1, \infty)$. Hence $M(\lambda)$ is decreasing on (0, 1). We now prove our main result for the density of short returns.

THEOREM 4.

$$\limsup_{n\to\infty} \frac{1}{n}|\log \mu(\mathcal{S}_n(\lambda))| = M(\lambda).$$

The lower bound was proven in [**4**]. It remains to prove the upper bound. In [**4**], the bound was proven under the assumption that $\psi(0)$ be less than one, which is essentially satisfied only for Bernoulli measures. Here we obtain the lower bound for all weakly ψ-mixing measures. Theorem 4 leads to the following corollary.

COROLLARY 5.

$$\limsup_{n\to\infty} \frac{1}{n}|\log \mu(\mathcal{C}_n(\delta))| = M(\delta).$$

Proof. Clearly, $\mathcal{C}_n(\delta) \subset \bigcup_{j=1}^{[\delta n]} B_n(j)$, which implies that $\mathcal{C}_n(\delta) \subset \bigcup_{0<\lambda\le\delta} \mathcal{S}_n(\lambda)$. In fact, the union consists of no more than n distinct sets. Hence

$$\mu(\mathcal{C}_n(\delta)) \le n \max_{0<\lambda\le\delta} \mu(\mathcal{S}_n(\lambda)),$$

which implies that $\limsup_{n\to\infty}(1/n)|\log \mu(\mathcal{S}_n(\lambda))| \le \min_{0<\lambda\le\delta} M(\lambda)$. The upper bound follows from the fact that $\mathcal{S}_n(\lambda) \subset \mathcal{C}_n(\delta)$ for every $\lambda \le \delta$. The statement now follows because M is monotonically decreasing on (0, 1). □

PROPOSITION 6.

$$\limsup_{n\to\infty} \frac{1}{n}|\log \mu(\mathcal{S}_n(\lambda))| \le M(\lambda).$$

Let us first prove the following inequality, which is of some interest in its own right.

LEMMA 7. *Let $\gamma \in (0, 1)$. Then for all $\lambda \in (0, 1)$ and all large enough n,*

$$\mu(B_n(j)) \ge e^{\mathcal{O}(n^\gamma)} Z_r(w) Z_{j-r}(w - 1),$$

where $j = [\lambda n]$ and $n = wj + r$ with $0 \le r < j$, $w = [n/j]$.

Proof. Using the implicit short-term periodicity of points in the set $B_n(j)$, we will estimate its measure from below by a product involving suitable partition functions Z_*. We treat the two cases (A) $\lambda \in (0, 1/2]$ and (B) $\lambda \in (1/2, 1)$ separately.

(A) Let us first deal with the case $0 < \lambda \leq 1/2$. Put $j = [\lambda n]$ and $w = [n/j]$. Then $n = wj + r$, where $r < j$ ($r = 0$ if $\lambda n \in \mathbb{N}$ and $1/\lambda \in \mathbb{N}$). For an n-cylinder $A_n \subset B_n(j)$, one has the decomposition

$$A_n = \left(\bigcap_{k=0}^{w-1} T^{-jk} A_j(A_n)\right) \cap T^{-wj} A_r(A_n),$$

where $w = [n/j]$ and $A_j(A_n)$ is the j-cylinder that contains the n-cylinder A_n. Let $\Delta \geq \Delta_0$ be such that $\Delta < r,\ j - r$, and put

$$\tilde{A}_n = \left(\bigcap_{k=0}^{w} T^{-jk} A_{r-\Delta}(A_n)\right) \cap \left(\bigcap_{k=0}^{w} T^{-jk-r} A_{j-r-\Delta}(A_n)\right).$$

Here we have opened up gaps of length Δ at each occurrence of period after length j, and then cut each period of length j into two pieces of lengths r and $r - j$. Since $A_n \subset \tilde{A}_n$, we clearly have $\mu(\tilde{A}_n) \geq \mu(A_n)$, and in order to get a comparison in the opposite direction, we let $\beta > 1$ and put

$$\mathcal{G}_{n,j} = \{A_n \in \mathcal{A}^n \mid A_n \subset B_n(j),\ \mu(A_n) \geq e^{-2w\Delta^\beta} \mu(\tilde{A}_n)\}$$

for the 'good' n-cylinders in $B_n(j)$ whose measures are comparable to the measure of $\tilde{A}_n$. Put $G_{n,j} = \bigcup_{A_n \in \mathcal{G}_{n,j}} A_n$. Then for every $A_n \subset B_n(j)$, one has

$$\mu\left(\bigcup_{A'_n \subset \tilde{A}_n \cap B_n(j), A'_n \in \mathcal{A}^n} A'_n\right) \geq (1 - |\mathcal{A}|^{2w\Delta} e^{-2w\Delta^\beta}) \mu(\tilde{A}_n \cap B_n(j)),$$

as $\tilde{A}_n \cap B_n(j) = \bigcup_{A'_n \subset \tilde{A}_n, A'_n \in \mathcal{G}_{n,j}} A'_n$. This implies that if $|\mathcal{A}|^{2w\Delta} e^{-2w\Delta^\beta} < 1$, then $\tilde{A}_n \cap B_n(j) \neq \emptyset$ if and only if there exists an $A'_n \subset \tilde{A}_n$, $A'_n \in \mathcal{A}^n$, which also belongs to $\mathcal{G}_{n,j}$. Hence

$$\mu(B_n(j)) \geq \mu(G_{n,j}) \geq e^{-2w\Delta^\beta} \sum_{\tilde{A}_n} \mu(\tilde{A}_n),$$

where the sum is over all $\tilde{A}_n$ for which there is an $A_n \subset B_n(j)$. Since all $A_n \subset B_n(j)$ are of the form $(x_1 \ldots x_j)^w x_1 \ldots x_r$, where $x_1 \ldots x_j$ runs through all possible periodic words of lengths j, we get

$$\begin{aligned}\sum_{\tilde{A}_n} \mu(\tilde{A}_n) &= (1 + \mathcal{O}(\psi^\pm(\Delta)))^{2w+1} \\ &\quad \times \sum_{x_1 \ldots x_{r-\Delta}} \sum_{x_{r+1} \ldots x_{j-\Delta}} \mu(A_{r-\Delta}(x_1 \ldots x_{r-\Delta}))^{w+1} \mu(A_{j-r-\Delta}(x_{r+1} \ldots x_{j-\Delta}))^w \\ &= (1 + \mathcal{O}(\psi^\pm(\Delta)))^{2w+1} Z_{r-\Delta}(w) Z_{j-r-\Delta}(w-1),\end{aligned}$$

where the sum is over all $(r - \Delta)$-words $x_1 \ldots x_{r-\Delta}$ and all $(j - n - \Delta)$-words $x_{r+1} \ldots x_{j-\Delta}$, with $Z_m(k) = \sum_{A_m \in \mathcal{A}^m} \mu(A_m)^{k+1}$. Hence

$$\mu(B_n(j)) \geq c_1 e^{-2w\Delta^\beta} Z_{r-\Delta}(w) Z_{j-r-\Delta}(w-1)$$

for some $c_1 > 0$. We have to choose $\Delta \geq \Delta_0$ and need to have $|\mathcal{A}|^\Delta e^{-\Delta^\beta} < 1$. This requires β to be bigger than one.

Next, we compare $Z_{r-\Delta}(w)$ to $Z_r(w)$ as follows:

$$Z_{r-\Delta}(w) = \sum_{x_1 \ldots x_{r-\Delta}} \mu(A_{r-\Delta}(x_1 \ldots x_{r-\Delta}))^{w+1} \geq \frac{1}{|\mathcal{A}|^\Delta} \sum_{x_1 \ldots x_r} \mu(A_{r-\Delta}(x_1 \ldots x_r))^{w+1},$$

as $\mu(A_r(x_1 \ldots x_r)) \leq \mu(A_{r-\Delta}(x_1 \ldots x_{r-\Delta}))$ and $\#\{A_r \mid A_r \subset A_{r-\Delta}\} \leq |\mathcal{A}|^\Delta$. Hence $Z_{r-\Delta}(w) \geq |\mathcal{A}|^{-\Delta} Z_r(w)$ and, similarly, $Z_{j-r-\Delta}(w-1) \geq |\mathcal{A}|^{-\Delta} Z_{j-r}(w-1)$. This implies that

$$\mu(B_n(j)) \geq c_1 |\mathcal{A}|^{-2\Delta} e^{-2w\Delta^\beta} Z_r(w) Z_{j-r}(w-1) \geq c_1 e^{-c_2 \Delta^{\alpha\beta}} Z_r(w) Z_{j-r}(w-1)$$

if we choose $\Delta = [j^\alpha]$ for some $\alpha \in (0, 1)$. If α is small enough, then $\gamma \geq \alpha\beta$.

(B) Now consider the case where $\lambda \in (1/2, 1)$. Again, we put $j = [n\lambda]$ and $n = j + r$ (note that $[1/\lambda] = 1$). If $A_n \subset B_n(j)$ is an n-cylinder, then $A_n = A_r(A_n) \cap T^{-j} A_r(A_n) \cap T^{-r} A_{n-2r}(T^r A_n)$, where $n - 2r \geq 0$. Let $\Delta \geq \Delta_0$ (not too large) and define, as above,

$$\tilde{A}_n = A_{r-\Delta}(A_n) \cap T^{-j} A_{r-\Delta}(A_n) \cap T^{-r} A_{n-2r-\Delta}(T^r A_n)$$

(if $n - 2r - \Delta > 0$; otherwise, just put $\tilde{A}_n = A_{r-\Delta}(A_n) \cap T^{-j} A_{r-\Delta}(A_n)$). For $\beta > 1$, we introduce as before the 'good set'

$$\mathcal{G}_{n,j} = \{A_n \in \mathcal{A}^n \mid A_n \subset B_n(j),\ \mu(A_n) \geq e^{-\Delta^\beta} \mu(\tilde{A}_n)\}.$$

If $|\mathcal{A}|^{2\Delta} e^{-\Delta^\beta} < 1$, then for every $\tilde{A}_n$ (of the form given above) there exists an $A_n \in \mathcal{G}_{n,j}$ such that $A_n \subset \tilde{A}_n$, and therefore

$$\mu(B_n(j)) \geq \mu(G_{n,j}) \geq e^{-\Delta^\beta} \sum_{\tilde{A}_n} \mu(\tilde{A}_n),$$

where the sum is over all $\tilde{A}_n = A_{r-\Delta}(x_1 \ldots x_{r-\Delta}) \cap T^{-j} A_{r-\Delta}(x_1 \ldots x_{r-\Delta}) \cap T^{-r} A_{r-\Delta}(x_{r+1} \ldots x_{n-j})$ (in the case where $n - 2r - \Delta > 0$) and $x_1 \ldots x_{r-\Delta}$, $x_{r+1} \ldots x_{n-j}$ are arbitrary words. Hence

$$\begin{aligned}
\mu(B_n(j)) &\geq (1 + \mathcal{O}(\psi^-(\Delta)))^2 e^{-\Delta^\beta} \sum_{x_1 \ldots x_{r-\Delta}} \mu(A_{r-\Delta}(x_1 \ldots x_{r-\Delta}))^2 \\
&\quad \times \sum_{x_{r+1} \ldots x_{j-\Delta}} \mu(A_{j-r-\Delta}(x_{r+1} \ldots x_{j-\Delta})) \\
&= (1 + \mathcal{O}(\psi^-(\Delta)))^2 e^{-\Delta^\beta} Z_{r-\Delta}(1) Z_{j-r-\Delta}(0) \\
&= (1 + \mathcal{O}(\psi^-(\Delta)))^2 |\mathcal{A}|^{-2\Delta} e^{-\Delta^\beta} Z_r(1) Z_{j-r}(0),
\end{aligned}$$

where in the last line we have used the comparison from the end of part (A). Again, we choose $\Delta = [j^\alpha]$, where $\alpha \in (0, 1)$ can be chosen small enough so that $\gamma \geq \alpha\beta$. □

Proof of Proposition 6. Obviously, $\mu(\mathcal{S}_n(\lambda)) \geq B_n(j)$ and therefore, by Lemma 7,

$$\frac{\log \mu(\mathcal{S}_n(\lambda))}{n} \geq -\frac{\mathcal{O}(n^\gamma)}{n} + \frac{1}{n} \log Z_r(w) + \frac{1}{n} \log Z_{j-r}(w-1),$$

where $\gamma < 1$ can be chosen arbitrarily. As $n \to \infty$, the first term goes to zero. Thus

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu(\mathcal{S}_n(\lambda)) \geq \liminf_{n\to\infty} \frac{1}{n} \log Z_r(w) + \liminf_{n\to\infty} \frac{1}{n} \log Z_{j-r}(w-1).$$

Now, notice that (as $n = w[\lambda n] + r$)

$$\frac{1}{n} \log Z_r(w) = \frac{r}{n}\frac{1}{r} \log Z_r(w) \to (1-\lambda\ell)\, W(\ell),$$

since

$$\frac{r}{n} = \frac{n-[\lambda n]w}{n} = 1 - \frac{[\lambda n]}{n}w \to 1-\lambda\ell \quad \text{and} \quad w \to \ell = [1/\lambda]$$

as $n \to \infty$ and W is continuous by Theorem 1. Similarly,

$$\frac{1}{n} \log Z_{j-r}(w) = \frac{j-r}{n}\frac{1}{j-r} \log Z_{j-r}(w) \to (\lambda(1+\ell)-1)W(\ell-1),$$

since

$$\frac{j-r}{n} = \frac{[n\lambda]-(n-[\lambda n])w}{n} = \frac{[n\lambda]}{n}(1+w) - 1 \to \lambda\,(1+\ell) - 1.$$

This implies the statement of the proposition. □

LEMMA 8.

$$\limsup_{n\to\infty} \frac{1}{n}|\log \mu(\mathcal{S}_n(\lambda))| \geq M(\lambda).$$

Proof. Again we treat the two cases (A) $\lambda \in (0, 1/2]$ and (B) $\lambda \in (1/2, 1)$ separately.

(A) $0 < \lambda \leq \frac{1}{2}$. As above, we write $n = wj + r$, where $j = [\lambda n]$ and $w = [n/j]$, with $0 \leq r < j$. Since all $A_n \subset B_n(j)$ are of the form $(x_1 \ldots x_j)^w x_1 \ldots x_r$, where $x_1 \ldots x_j$ runs through all possible periodic words of lengths j, we get (by summing over such A_n) that

$$\sum_{A_n} \mu(A_n) \leq (1+\psi^+(0))^{2w+1} \sum_{x_1\ldots x_r} \sum_{x_{r+1}\ldots x_j} \mu(A_r(x_1\ldots x_r))^{w+1} \mu(A_{j-r}(x_{r+1}\ldots x_j))^w,$$

where the sum is over all r-words $x_1 \ldots x_r$ and all $(j-n)$-words $x_{r+1} \ldots x_j$. Hence

$$\mu(\mathcal{S}_n(\lambda)) \leq (1+\psi^+(0))^{2w+1} Z_r(w) Z_{j-r}(w-1)$$

and therefore, as in the proof of Proposition 6,

$$\lim_{n\to\infty} \frac{1}{n}|\log \mu(\mathcal{S}_n(\lambda))| \geq \lim_{n\to\infty} \frac{1}{n}|\log Z_r(w)| + \lim_{n\to\infty} \frac{1}{n}|\log Z_r(w-1)| = M(\lambda).$$

(B) $\lambda \in (\frac{1}{2}, 1)$. We again put $j = [n\lambda]$ and $n = j + r$, with $0 \leq r < j$ (as $[1/\lambda] = 1$). If $A_n \subset B_n(j)$ is an n-cylinder, then $A_n = A_r(A_n) \cap T^{-j} A_r(A_n) \cap T^{-r} A_{n-2r}(T^r A_n)$, where $n - 2r \geq 0$. Hence

$$\mu(B_n(j)) \leq (1+\psi^+(0))^2 \sum_{x_1\ldots x_r} \mu(A_r(x_1\ldots x_r))^2 \sum_{x_{r+1}\ldots x_j} \mu(A_{j-r}(x_{r+1}\ldots x_j))$$

for arbitrary words $x_1 \ldots x_r, x_{r+1} \ldots x_{n-j}$. This implies that

$$\mu(\mathcal{S}_n(\lambda)) \leq c_1 Z_r(1) Z_{j-r}(0) \; (c_1 > 0)$$

and

$$\lim_{n\to\infty} \frac{1}{n}|\log \mu(\mathcal{S}_n(\lambda))| \geq \lim_{n\to\infty} \frac{1}{n}|\log Z_r(1)| = M(\lambda).$$ □

Proof of Theorem 4. The theorem now follows from Proposition 6 and Lemma 8. □

Besides the exact limiting behavior we get from Theorem 4, we can also prove the following simpler bounds.

LEMMA 9.

$$\liminf_{n\to\infty} \frac{1}{n}|\log \mu(\mathcal{C}_n(\delta))| \le h_\mu(1-\delta)$$

for all $\delta \in (0, 1)$.

Proof. As before, let $j = [\delta n]$, and for an n-cylinder A_n in $B_n(j)$, we put $\tilde{A}_n = A_{r-\Delta}(A_n) \cap T^{-r} A_{n-r}(T^r A_n)$ for a gap of length Δ on the segment $[r - \Delta + 1, r]$. Let $\beta > 1$ and

$$\mathcal{G} = \{A_n \in \mathcal{A}^n \mid A_n \subset B_n(j),\ \mu(A_n) \ge e^{-\Delta^\beta} \mu(\tilde{A}_n)\},$$

and observe that if $|\mathcal{A}|^\Delta e^{-\Delta^\beta} < 1$, then

$$\tilde{A}_n \cap B_n(j) \ne \emptyset \iff \text{there exists } A'_n \in \mathcal{G} \text{ with } A'_n \subset \tilde{A}_n.$$

Hence if $\Delta \ge \Delta_0$, then

$$\begin{aligned}
\mu(B_n(j)) &\ge \mu(G) \\
&\ge e^{-\Delta^\beta} \sum_{\tilde{A}_n} \mu(\tilde{A}_n) \\
&\ge (1 - \psi^-(\Delta)) e^{-\Delta^\beta} \sum_{x_1 \ldots x_{j-\Delta}} \mu(A_{j-\Delta}(x_1 \ldots x_{j-\Delta})) \\
&\quad \times \sum_{x_{j-\Delta+1} \ldots x_j} \mu(A_{n-j}(T^j (x_1 \ldots x_j)^\infty)),
\end{aligned}$$

where $G = \bigcup_{A_n \in \mathcal{G}} A_n$, as $\tilde{A}_n = A_{j-\Delta}(x_1 \ldots x_{j-\Delta}) \cap A_{n-j}((x_1 \ldots x_j)^\infty)$. By the Shannon–McMillan–Breiman theorem [**30**], for every $\varepsilon > 0$ there exists a set $\Omega_\varepsilon \subset \Omega$ with measure at least $1 - \varepsilon$ so that $\mu(A_{n-j}((x_1 \ldots x_j)^\infty)) \ge \exp(-(n - j)(h_\mu + \varepsilon))$ for all sufficiently large n and all $(x_1 \ldots x_j)^\infty$ such that $A_{n-j}((x_1 \ldots x_j)^\infty) \cap \Omega_\varepsilon \ne \emptyset$. Hence

$$\begin{aligned}
\mu(\mathcal{C}_n(\delta)) &\ge e^{-\Delta^\beta} (1 - \psi^-(\Delta)) \\
&\quad \times \sum_{A_n \in \mathcal{G}, T^j A_n \cap \Omega_\varepsilon \ne \emptyset} \mu(A_{j-\Delta}(T^j A_n)) e^{-(n-j)(h_\mu+\varepsilon)} \\
&\ge e^{-\Delta^\beta} (1 - \psi^-(\Delta)) e^{-(n-j)(h_\mu+\varepsilon)} \\
&\quad \times \left(\sum_{x_1 \ldots x_{j-\Delta}} \mu(A_{j-\Delta}(x_1 \ldots x_{j-\Delta})) - \varepsilon \right) \\
&\ge e^{-\Delta^\beta - (n-j)(h_\mu+\varepsilon)} (1 - \psi^-(\Delta))\, (1 - \varepsilon)
\end{aligned}$$

and, consequently, $\lim_{n\to\infty} (1/n) \log \mu(\mathcal{C}_n(\delta)) \ge -(1 - \delta)(h_\mu + \varepsilon)$ if we take $\Delta = [n^\alpha]$, where $\alpha < 1$ is such that $\beta\alpha < 1$. Now let $\varepsilon \to 0^+$ to get the result. □

LEMMA 10. *Let* γ_μ *be as in Theorem 1. Then*

$$\liminf_{n\to\infty} \frac{1}{n}|\log \mu(\mathcal{C}_n(\delta))| \ge \gamma_\mu(1-\delta).$$

The proof is exactly the same as that of [**4**, Proposition 1(a)]. It uses the mixing properties of ψ-mixing measures without the assumption that $\psi(0) < 1$.

4. *Uniform decay rate of cylinders and the free energy*

In this section we compute the free energy $F(\beta)$ of the process τ_n; it is defined by

$$F(\beta) \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n}\log \int_\Omega \exp(\beta\tau_n(A_n))\,d\mu = \lim_{n\to\infty} \frac{1}{n}\log \sum_{j=1}^{\infty} e^{\beta j}\mathbb{P}(\tau_n = j) \tag{5}$$

whenever the limit exists (we use the probabilists' shorthand $\mathbb{P}(\tau_n = j)$ for $\mu(\{x \in \Omega \mid \tau_n(x) = j\})$).

THEOREM 11. *Let μ be a weakly ψ-mixing measure. Then*

$$F(\beta) = \begin{cases} \beta & \text{if } -\gamma_\mu \le \beta < 0, \\ -\gamma_\mu & \text{if } \beta \le -\gamma_\mu. \end{cases}$$

Remark 2. Although $F(\beta)$ is not differentiable, one may still take its Legendre transform $\mathcal{L}F(\delta)$ and produce an upper bound for the deviation function $M(\delta)$ (see [**19**]). We immediately get

$$M(\delta) \le \mathcal{L}F(\delta) = -\gamma_\mu(1-\delta),$$

which is consistent with the bound obtained in Lemma 10.

Remark 3. The proof of the theorem splits into two parts. The first part consists of obtaining an upper bound for the sum $\sum_{j=1}^{\infty} e^{\beta j}\mathbb{P}(\tau_n = j)$, which is achieved by using the mixing properties of the measure. We defer to the proof of this bound in [**4**, Proposition 6], which applies verbatim to our situation (it does not require the stringent condition $\psi(0) < 1$). This upper bound allows us to show that $\limsup_{n\to\infty}(1/n)\log \int_\Omega \exp(\beta\tau_n(A_n))\,d\mu$ is piecewise constant as prescribed in Theorem 11. However, the lower bound is more interesting. Here we need an additional property of our measure, namely the existence of a sequence of cylinders whose measures decay exponentially to zero at a rate which is exactly the constant γ_μ given by Theorem 1 and whose first return is sublinear. This sequence is explicitly constructed in Lemma 13 below. We will give the proof of the lower bound after proving Lemma 13.

As before, let $\gamma_\mu = \liminf_n (1/n)|\log b_n|$ be the exponential decay rate of the measures of n-cylinders, where $b_n = \max_{A_n\in\mathcal{A}^n} \mu(A_n)$ and $0 < |\log\eta| \le \gamma_\mu \le h_\mu$ by Lemma 3.

LEMMA 12. *There exists a sequence of n-cylinders A_n, $n = 1, 2, 3, \ldots,$ such that $\gamma_\mu = \lim_{n\to\infty}(1/n)|\log\mu(A_n)|$.*

Proof. We have to show that the lim inf is equal to the limit along a suitable sequence of cylinders. For this purpose, let A_{n_j} (with n_j an increasing sequence) be a sequence of n_j-cylinders such that $\gamma_\mu = \lim_j(1/n_j)|\log\mu(A_{n_j})|$. Let $\varepsilon > 0$, and take J large enough so that $|(1/n_j)|\log\mu(A_{n_j})| - \gamma_\mu| < \varepsilon/2$ for all $j \ge J$. Let $\alpha \in (0, 1)$ and $\Delta = [n_j^\alpha]$ ($\Delta \ge \Delta_0$ if n_j is big enough), and put $\tilde{A}_{n_j+(k-1)\Delta} = \bigcap_{i=0}^{k-1} T^{-i(n_j+\Delta)}A_{n_j}$; by Lemma 2, this implies that

$$\mu(\tilde{A}_{n_j+(k-1)\Delta}) = \mu(A_{n_j})^k(1 + \mathcal{O}(\psi^{\pm}(\Delta)))^{k-1}.$$

Now choose $\beta > 1$ so that $\alpha\beta < 1$, and proceed as before by defining

$$\mathcal{G}_k = \{A_{kn_j+(k-1)\Delta} \subset \tilde{A}_{kn_j+(k-1)\Delta} \mid \mu(A_{kn_j+(k-1)\Delta}) \ge e^{-k\Delta^\beta}\mu(\tilde{A}_{kn_j+(k-1)\Delta})\}.$$

Since $\beta > 1$, we have for all Δ large enough that $|\mathcal{A}|^{(k-1)\Delta} e^{-k\Delta^\beta} < 1$, which implies that there exists at least one cylinder $A_{kn_j+(k-1)\Delta} \subset \tilde{A}_{kn_j+(k-1)\Delta}$, $A_{kn_j+(k-1)\Delta} \in \mathcal{A}^{kn_j+(k-1)\Delta}$, so that $\mu(A_{kn_j+(k-1)\Delta}) \geq e^{-k\Delta^\beta} \mu(\tilde{A}_{kn_j+(k-1)\Delta})$; therefore

$$\begin{aligned}\frac{\log \mu(A_{kn_j+(k-1)\Delta})}{kn_j + (k-1)\Delta} &\geq -\frac{k\Delta^\beta}{kn_j + (k-1)\Delta} + \frac{k \log \mu(A_{n_j})}{kn_j + (k-1)\Delta} + \frac{k \log (1 - \psi^-(\Delta))}{kn_j + (k-1)\Delta} \\ &\geq -2\frac{\Delta^\beta}{n_j} + \frac{1}{1 + \Delta/n_j} \frac{\log \mu(A_{n_j})}{n_j} - 2\frac{\psi^-(\Delta)}{n_j} \\ &\geq -cn_j^{\beta\alpha-1} + \frac{\log \mu(A_{n_j})}{n_j},\end{aligned}$$

where we put $\Delta = [n_j^\alpha]$ and $c \approx 3 + 2\gamma_\mu$ (as $1/(1 + \Delta/n_j) \leq 1 + 2\Delta/n_j$ for j large enough). Hence

$$\left| \frac{|\log \mu(A_{kn_j+(k-1)\Delta})|}{kn_j + (k-1)\Delta} - \gamma_\mu \right| < \frac{\varepsilon}{2} + \frac{c}{n_j^{1-\beta\alpha}} < \varepsilon$$

for all k if n_j is large enough, as $\alpha\beta < 1$. □

LEMMA 13. *There exists a sequence of cylinders $B_j \in \mathcal{A}^j$ such that*

$$\lim_{j\to\infty} \frac{1}{j} |\log \mu(B_j)| = \gamma_\mu \quad \textit{and} \quad \lim_{j\to\infty} \frac{1}{j} \tau(B_j) = 0.$$

Proof. By Lemma 12, there exists a sequence of cylinders $A_n \in \mathcal{A}^n$ such that $(1/n)|\log \mu(A_n)| \to \gamma_\mu$ as $n \to \infty$. Let $\varepsilon > 0$, and take N so that

$$\left| \frac{1}{n} |\log \mu(A_n)| - \gamma_\mu \right| \leq \varepsilon/3 \quad \text{for all } n \geq N.$$

Let $\alpha, \alpha' \in (0, 1)$ and put $k_n = [n^{\alpha'}]$, $\Delta_n = [n^\alpha]$. Also write, for simplicity, $n' = n + \Delta_n$ and $(n+1)' = n + 1 + \Delta_{n+1}$. Then

$$k_{n+1}(n+1)' - \Delta_{n+1} - (k_n n' - \Delta_n) \in \begin{cases} [0, 3] & \text{if } k_{n+1} = k_n, \\ [k_n + \Delta_n, k_n + \Delta_n + 3] & \text{if } k_{n+1} = k_n + 1. \end{cases}$$

Let $\epsilon_n = k_{n+1} - k_n$ ($\epsilon_n = 0, 1$), and for $j \in \big[k_n n' - \Delta_n, k_{n+1}(n+1)' - \Delta_{n+1}\big)$ put

$$\mathcal{D}_j = \{D \in \mathcal{A}^{(k_n+\epsilon_n)n'-\Delta_n} \mid D \subset \tilde{A}_n\},$$

where

$$\tilde{D}_j = \bigcap_{j=0}^{k_n+\epsilon_n-1} T^{-jn'} A_n \in \bigvee_{j=0}^{k_n+\epsilon_n-1} T^{-jn'} \mathcal{A}^n.$$

For $\beta > 1$, we define the 'good' set of cylinders in $\tilde{\mathcal{A}}^j$ whose measures are comparable to the measure of $\tilde{D}_n$:

$$\mathcal{G}_j = \{D \in \mathcal{A}^j \mid \mu(D) \geq \exp(-(k_n + \epsilon_n)\Delta_n^\beta)\mu(\tilde{D}_j)\}.$$

If $|\mathcal{A}|^{\Delta} e^{-\Delta^{\beta}} < 1$, then $\mathcal{G}_j \neq \emptyset$. Hence we can find a j-cylinder $B_j \in \mathcal{G}_j$ such that $B_j \subset \tilde{D}_j$ which, moreover, has comparable measure, i.e. $\mu(B_j) \geq \exp(-(k_n + \epsilon_n)\Delta_n^{\beta})\mu(\tilde{D}_j)$. By the mixing property,

$$\mu(\tilde{D}_j) = (1 + \mathcal{O}(\psi^{\pm}(\Delta_n)))^{k_n+\epsilon_n-1}\mu(A_n)^{k_n+\epsilon_n},$$

which implies that

$$\log \mu(B_j) \geq -(k_n + \epsilon_n)\Delta_n^{\beta} + k_n \log \mu(A_n) + k_n \log\,(1 - \psi^{-}(\Delta_n)).$$

If $\alpha' + \beta\alpha < 1$ and n is large enough, then $(1/j)(k_n + \epsilon_n)\Delta_n^{\alpha} < \varepsilon/3$ and $(1/j)k_n \log(1 - \psi^{-}(\Delta_n)) < \varepsilon/3$. Hence $|(1/j)|\log \mu(B_j)| - \gamma_\mu| < \varepsilon$ for all sufficiently large j. Moreover, we note that $\tau(A'_j) \leq n + \Delta_n$, which implies that $\lim_{j\to\infty}(1/j)\tau(B_j) = 0$. $\square$

Proof of Theorem 11. As described in Remark 3, it will be sufficient to show that

$$\liminf_{n\to\infty} \frac{1}{n} \log \int_{\Omega} \exp(\beta\tau_n(A_n))\,d\mu \geq \begin{cases} \beta & \text{if } -\gamma_\mu \leq \beta < 0, \\ -\gamma_\mu & \text{if } \beta \leq -\gamma_\mu. \end{cases}$$

We have two cases.

(i) $-\gamma_\mu \leq \beta < 0$. The result immediately follows, since

$$\sum_{j=1}^{\infty} \exp(\beta j)\mathbb{P}(\tau_n = j) \geq \exp(\beta n + \Delta),$$

where Δ was introduced at the beginning of §3.

(ii) $\beta < -\gamma_\mu$. In any partition $\mathcal{A}^n$, let us choose a cylinder A'_n which satisfies Lemma 13. Then

$$\sum_{j=1}^{\infty} e^{\beta j}\mathbb{P}(\tau_n = j) \geq \exp(\beta\tau_n(A'n))\mu(A'_n).$$

But $\mu(A'_n)$ decays exponentially fast to zero with a rate given by $-\gamma_\mu$, while $(1/n)\tau_n(A'n)$ goes to zero. This concludes the proof. $\square$

Acknowledgements. We gladly acknowledge numerous fruitful discussions with M. Abadi about the Rényi entropy and its applications. We also thank J.-R. Chazottes for bringing [**29**] to our attention.

This work was supported by the NSF (grant DMS-0301910) and by the CNRS unité FR2291 FRUMAN (#64636).

REFERENCES

[1] M. Abadi. Sharp error terms and necessary conditions for exponential hitting times in mixing processes. *Ann. Probab.* **32** (2004), 243–264.

[2] M. Abadi. Hitting, returning and the short correlation function. *Bull. Braz. Math. Soc.* **37** (2006), 593–609.

[3] M. Abadi and A. Galves. Inequalities for the occurrence times of rare events in mixing processes. The state of the art. *Markov Process. Related Fields.* **7** (2001), 97–112.

[4] M. Abadi and S. Vaienti. Large deviations for short recurrence. *Discrete. Contin. Dyn. Syst.* **21**(3) (2008), 729–747.

[5] V. Afraimovich. Pesin's dimension for Poincaré recurrence. *Chaos* **7**(1) (1997), 11–20.

[6] V. Afraimovich, J.-R. Chazottes and B. Saussol. Pointwise dimensions for Poincaré recurrence associated with maps and special flows. *Discrete Contin. Dyn. Syst.* **9** (2003), 263–280.

[7] V. Afraimovich, E. Ugalde and J. Urias. *Fractal Dimensions for Poincaré Recurrence (Monograph Series on Nonlinear Sciences and Complexity, 2)*. Elsevier, Amsterdam, 2006.

[8] L. Barreira, Ya. Pesin and J. Schmeling. On a general concept of multifractality: multifractal spectra for dimensions, entropies and Lyapunov exponents. *Chaos* **7**(1) (1997), 27–38.

[9] C. Beck and F. Schlögl. *Thermodynamics of Chaotic Systems*. Cambridge University Press, Cambridge, 1993.

[10] D. Bessis, G. Paladin, G. Turchetti and S. Vaienti. Generalized dimensions, entropies and Lyapunov exponents from the pressure function for strange sets. *J. Stat. Phys.* **51** (1988), 109–134.

[11] P. Billingsley. *Convergence in Probability Measures*. Wiley, New York, 1968.

[12] C. Bonanno, S. Galatolo and S. Isola. Recurrence and algorithmic information. *Nonlinearity* **17** (2004), 1057–1074.

[13] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Mathematics, 470)*. Springer, New York, 1975.

[14] J.-R. Chazottes and F. Durand. Local rates of Poincaré recurrence for rotations and weak-mixing. *Discrete Contin. Dyn. Syst.* **12** (2005), 175–183.

[15] P. Collet, A. Galves and B. Schmitt. Fluctuations of repetition times for Gibbsian sources. *Nonlinearity* **12** (1999), 1225–1237.

[16] A. Csordas and P. Szepfalusy. Generalized entropy decay rate of one-dimensional maps. *Phys. Rev.* A **38** (1989), 2582–2587.

[17] P. Doukhan. *Mixing Properties and Examples (Lecture Notes in Statistics, 85)*. Springer, Berlin, 1995.

[18] J.-P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57** (1985), 617–656.

[19] R. S. Ellis. *Entropy, Large Deviations and Statistical Mechanics*. Springer, Berlin, 1985.

[20] P. Ferrero, N. Haydn and S. Vaienti. Entropy fluctuations for parabolic maps. *Nonlinearity* **16** (2003), 1203–1218.

[21] A. Galves and B. Schmitt. Inequalities for hitting times in mixing dynamical systems. *Random Comput. Dyn.* **5** (1997), 337–348.

[22] P. Grassberger and I. Procaccia. Estimation of the Kolmogorov entropy from a chaotic signal. *Phys. Rev.* A **28** (1983), 2591–2593.

[23] P. Grassberger and I. Procaccia. Dimensions and entropies of strange attractors from a fluctuating dynamics approach. *Phys.* D **13** (1984), 34–54.

[24] N. Haydn, Y. Lacroix and S. Vaienti. Hitting and return times in ergodic dynamical systems. *Ann. Probab.* **33** (2005), 2043–2050.

[25] N. Haydn and S. Vaienti. The limiting distribution and error terms for return time of dynamical systems. *Discrete Contin. Dyn. Syst.* **10** (2004), 584–616.

[26] M. Hirata, B. Saussol and S. Vaienti. Statistics of return times: a general framework and new applications. *Comm. Math. Phys.* **206** (1999), 33–55.

[27] M. Kupsa. Local return rates in Sturmian sub-shifts. *Acta Univ. Carolin. Math. Phys.* **44** (2003), 17–28.

[28] P. Kurka. Local return rates in substitutive sub-shifts. *Acta Univ. Carolin. Math. Phys.* **44** (2003), 29–42.

[29] T. Luczak and W. Szpankowski. A suboptimal lossy data compression based on approximate pattern matching. *IEEE Trans. Inform. Theory* **43** (1997), 1439–1451.

[30] R. Mañé. *Ergodic Theory and Differentiable Dynamics*. Springer, Berlin, 1987.

[31] G. Paladin, G. Parisi and A. Vulpiani. Intermittency in chaotic systems and Rényi entropies. *J. Phys. A* **19** (1986), L997–L1001.

[32] V. Penné, B. Saussol and S. Vaienti. Dimensions for recurrence times: topological and dynamical properties. *Discrete Contin. Dyn. Syst.* **5** (1999), 783–798.

[33] Ya. Pesin. *Dimension Theory in Dynamical Systems*. University of Chicago Press, Chicago, 1997.

[34] A. Rényi. *Probability Theory*. North Holland, Amsterdam, 1970.

[35] B. Saussol, S. Troubetzkoy and S. Vaienti. Recurrence, dimensions and Lyapunov exponents. *J. Stat. Phys.* **106** (2002), 623–634.

[36] F. Takens and E. Verbitsky. Generalized entropies: Rényi and correlation integral approach. *Nonlinearity* **11**(4) (1998), 771–782.

[37] F. Takens and E. Verbitsky. Multifractal analysis of local entropies for expansive homeomorphisms with specification. *Comm. Math. Phys.* **203** (1999), 593–612.