A note on large deviations for unbounded observables *

Matthew Nicol † Andrew Török ‡§

October 27, 2019

Contents

Abstract

We consider exponential large deviations estimates for unbounded observables on uniformly expanding dynamical systems. We show that uniform expansion does not imply the existence of a rate function for unbounded observables no matter the tail behavior of the cumulative distribution function. We give examples of unbounded observables with exponential decay of autocorrelations, exponential decay under the transfer operator in each L^p, $1 \leq p < \infty$, and strictly stretched exponential large

---

*The authors would like to thank Vaughn Climenhaga for helpful discussions.

†Department of Mathematics, University of Houston, Houston, USA. e-mail: ⟨nicol@math.uh.edu⟩. MN thanks the NSF for partial support on NSF-DMS Grant 1600780.

‡Department of Mathematics, University of Houston, Houston Texas, USA. e-mail: ⟨torok@math.uh.edu⟩. AT thanks the NSF for partial support on NSF-DMS Grant 1816315.

§Institute of Mathematics of the Romanian Academy, Bucharest, Romania.

deviation. For observables of form $|\log d(x,p)|^{\alpha}$, p periodic, on uniformly expanding systems we give the precise stretched exponential decay rate. We also show that a classical example in the literature of a bounded observable with exponential decay of autocorrelations yet with no rate function is degenerate as the observable is a coboundary.

1 Introduction and brief background.

Suppose (T, X, μ) is a probability preserving transformation and $\varphi : X \to \mathbb{R}$ is a mean-zero integrable function i.e. $\mathbb{E}(\varphi) := \int_X \varphi \, d\mu = 0$. Throughout this paper we will write $S_n(\varphi) := \varphi + \varphi \circ T + \ldots + \varphi \circ T^{n-1}$ for the nth ergodic sum of φ. Sometimes we will write S_n instead of $S_n(\varphi)$ for simplicity of notation or when φ is clear from context. If we put $\overline{\varphi} = \int \varphi d\mu$ then ergodicity implies $\lim_{n\to\infty} \frac{1}{n} S_n(\varphi) = \overline{\varphi}$.

Large deviations theory concerns the rate of convergence of $\frac{1}{n} S_n(\varphi)$ to $\overline{\varphi}$.

If for each $\varepsilon > 0$ there exists $C > 0$ and $0 < \theta < 1$ such that for all $n \geq 0$

$$\mu(|\frac{1}{n} S_n(\varphi) - \overline{\varphi}| > \varepsilon) \leq C\theta^n$$

we will say $S_n(\varphi)$ has *large deviations with an exponential rate.*

If for each $\varepsilon > 0$ there exists $C > 0$ and $0 < \gamma < 1$ such that for all $n \geq 0$

$$\mu(|\frac{1}{n} S_n(\varphi) - \overline{\varphi}| > \varepsilon) \leq Ce^{-n^{\gamma}}$$

we say $S_n(\varphi)$ has *large deviations with a stretched exponential rate of order* $\gamma > 0$.

We now recall the definition of rate function and some other notions of large deviations theory pertaining to our results.

Definition 1.1. *A mean-zero integrable function $\varphi : \Omega \to \mathbb{R}$ is said to satisfy a large deviation principle with rate function $I(\alpha)$, if there exists a non-empty neighborhood U of 0 and a strictly convex function $I : U \to \mathbb{R}$, non-negative and vanishing only at $\alpha = 0$, such that*

$$\lim_{n\to\infty} \frac{1}{n} \log \mu(S_n(\varphi) \geq n\alpha) \quad = \quad -I(\alpha) \tag{1.1}$$

for all $\alpha > 0$ in U and

$$\lim_{n\to\infty} \frac{1}{n} \log \mu(S_n(\varphi) \leq n\alpha) \quad = \quad -I(\alpha) \tag{1.2}$$

for all $\alpha < 0$ in U.

In the literature this is referred to as a first level or local (near the average) large deviations principle.

If φ is a mean-zero continuous observable on an SRB attractor then φ has exponential large deviations (see [You90, Theorem 2 (2)]). For mean-zero Hölder observables on Young Towers with exponential tails (we refer to [You98] or [MN08, Sections 2 and 4] for the definition) which are not L^1 coboundaries in the sense that $\varphi \neq \psi \circ T - \psi$ for any $\psi \in L^1(\mu)$) such an exponential large deviations result holds with rate function $I_{\varphi}(\alpha)$ [MN08, RBY08]. A formula for the width of U is given in [RBY08] following a standard approach but it is not useful in concrete estimates.

2 Erdős-Rényi laws: background.

Erdős-Rényi laws [ER70] give estimates on the size of time-windows over which we should expect to average to achieve nontrivial almost sure limit laws.

Proposition 2.1 below is found in a proof from Erdős and Rényi [ER70] (see [CR79, Theorem 2.4.3], Grigull [Gri93], Denker and Kabluchko [DK07] or [DN13] where this method has been used). The Gauss bracket [.] denotes the integer part of a number. Throughout the proofs of this paper we will concentrate on the case $\alpha > 0$ as the case $\alpha < 0$ is identical with the obvious modifications of statements.

Proposition 2.1 ([ER70, DN13]). *(a) Suppose that φ satisfies a large deviation principle with rate function I defined on the open set U. Let $\alpha > 0$, $\alpha \in U$ and set*

$$\ell_n = \ell_n(\alpha) = \left[\frac{\log n}{I(\alpha)}\right] \qquad n \in \mathbb{N}.$$

Then the upper Erdős-Rényi law holds, that is, for μ a.e. $x \in X$

$$\limsup_{n\to\infty} \max_{0\le j\le n-\ell_n} \frac{1}{\ell_n} S_{\ell_n}(\varphi)\circ T^j(x) \le \alpha.$$

(b) If for some constant $C > 0$ and integer $\tau \ge 0$ for each interval A

$$\mu\left(\bigcap_{m=0}^{n-\ell_n}\{S_{\ell_n}(\varphi)\circ T^m \in A\}\right) \;\le\; C[\mu(S_{\ell_n}\in A)]^{n/(\ell_n)^\tau} \tag{2.1}$$

then the lower Erdős-Rényi law holds as well, that is, for μ a.e. $x \in X$

$$\liminf_{n\to\infty} \max_{0\le j\le n-\ell_n} \frac{1}{\ell_n} S_{\ell_n}(\varphi)\circ T^j \ge \alpha.$$

Remark 2.2. *If both Assumptions (a) and (b) of Proposition 2.1 hold then*

$$\lim_{n\to\infty} \max_{0\le m\le n-\ell_n} \frac{S_{\ell_n}\circ T^m}{\ell_n} = \alpha.$$

Remark 2.3. *Note that regularity of φ is not needed for the proof of Proposition 2.1 and that Proposition 2.1 (a) applies to unbounded observables.*

3 Rate functions and unbounded observables.

Alves et al [AFLV11] have shown that if T is a $C^{1+\delta}$ local diffeomorphism of an interval which has stretched exponential decay of correlations for Hölder functions versus bounded functions then there exists a Young Tower for T with stretched exponential tails. It is unknown whether *exponential* decay of correlations for Hölder functions versus bounded functions implies there exists a Young Tower for T with *exponential* tails in the same setting.

The proofs of [MN08, RBY08] of exponential large deviations with a rate function for a Hölder function φ on a dynamical system modeled by a Young Tower with exponential tails use spectral techniques. It is necessary to establish the analyticity of the linear operator $P_z : \mathcal{B} \to \mathcal{B}$, $z \in \mathbb{C}$, defined by $P_z v = P(e^{z\varphi} v)$ where P is a transfer operator with spectral gap on a Banach space of functions $\mathcal{B}$. This method of proof fails if φ is unbounded. We remark that recently the almost sure invariance principle has been proved for certain classes of unbounded functions on both uniformly expanding maps and intermittent type maps with a neutral fixed point [DGM10]. In this section we give an example to show that we cannot expect exponential large deviations with a rate function for unbounded observables on uniformly expanding maps, even if they satisfy the monotonicity and moment conditions of [DGM10] and the tails of the cumulative distribution function $P(\varphi > t)$ decay at any prescribed rate.

Theorem 3.1. *Suppose $T : X \to X$ is a measure preserving map of a compact Riemannian manifold (X, μ). Let p be a periodic point of period τ. Suppose that T is C^1 on an open neighborhood of the orbit of p. Suppose also there exists $\gamma > 0$ such that $T^n(x) \in B_{n^{-\gamma}}(p)$ i.o. for μ a.e. $x \in X$. Let φ be a continuous observable on $X/\{p\}$ such that $\lim_{x\to p} \varphi(p) = \infty$, $\int \varphi d\mu = 0$ and $\varphi > -\rho$ for some $\rho > 0$. Then the stationary stochastic process $\{\varphi \circ T^j\}$ does not satisfy exponential large deviations with a rate function.*

Proof. Without loss of generality we take the period of p to be one. If φ satisfies a large deviation principle with rate function I defined on an open set U then by Proposition 2.1 (a) if $\alpha > 0$ is in the interval where $I(\alpha)$ is defined and we let

$$\ell_n = \ell_n(\alpha) = \left[\frac{\log n}{I(\alpha)}\right] \qquad n \in \mathbb{N}$$

the upper Erdős-Rényi law holds, that is, for μ a.e. $x \in X$

$$\limsup_{n\to\infty} \max\{S_{\ell_n}(\varphi) \circ T^j(x)/\ell_n : 0 \le j \le n - \ell_n\} \le \alpha.$$

Since T is C^1, $|DT|_\infty < K$ for some $K > 0$ on an open neighborhood of the orbit of p.

Fix $\alpha > 0$ in U and let $M > \left[\frac{\alpha+\rho}{I(\alpha)}\right] \frac{2\log K}{\gamma}$. Choose N large enough that $\varphi(x) > M$ for all x such that $d(x,p) < \frac{1}{N^{\gamma/2}}$.

If $d(T^n(x), p) \le \frac{1}{n^\gamma}$ then $d(T^{n+j}(x), p) \le \frac{1}{n^{\gamma/2}}$ for at least j iterates, $1 \le j \le \frac{\gamma \log n}{2\log K}$ (this estimate comes from solving $K^j \frac{1}{n^\gamma} = \frac{1}{n^{\gamma/2}}$). Moreover if $n > N$, $j \le \frac{\gamma\log n}{2\log K}$ and $T^{n+j}(x) \in [0, \frac{1}{n^\gamma}]$, then $\varphi(T^{n+j}(x)) \ge M$. By assumption $T^n(x) \in B_{n^{-\gamma}(p)}$ i.o. for μ a.e. $x \in X$. If $T^n(x) \in B_{n^{-\gamma}}(p)$ then $S_{\ell_n}(\varphi) \circ T^j(x) > M(\gamma\frac{\log n}{2\log K}) - \rho\frac{\log n}{I(\alpha)}$ (as $\varphi \ge -\rho$). Since $M > \left[\frac{\alpha+\rho}{I(\alpha)}\right] \frac{2\log K}{\gamma}$ this implies that for μ a.e. x

$$\limsup_{n\to\infty} \max\{S_{\ell_n}(\varphi) \circ T^j(x)/\ell_n : 0 \le j \le n - \ell_n\} > \alpha$$

which is a contradiction to the upper Erdős-Rényi law. Hence exponential large deviations with a rate function cannot hold for this observable. ☐

Moreover, for such observables the logarithmic moment generating function is infinite, as shown by the next proposition.

Proposition 3.2. *Suppose $T : X \to X$ is a measure preserving map of a compact manifold (X, μ). Let p be a periodic point such that on an open neighborhood of the orbit of p the map T is C^1, and there are positive constants d and c such that $\mu(\{x : d(x,p) < r\}) \geq cr^d$ for small $r > 0$. Let φ be an observable on X such that $\lim_{x\to p} \varphi(p) = \infty$, $\int \varphi d\mu = 0$.*

Then for any $t > 0$

$$\lim_{n\to\infty} \frac{1}{n} \log \left(\int e^{tS_n(\varphi)} d\mu \right) = \infty \tag{3.1}$$

Proof. Without loss of generality, can assume that p is a fixed point.

Pick $\lambda > 1$ such that $|T'| < \lambda$ in a neighborhood of p.

Fix $t > 0$; for $L > 0$, let $M > L/t$ and $r > 0$ such that

$$d(x,p) < r \implies \varphi(x) > M.$$

Since

$$d(x,p) < \frac{r}{\lambda^n} \implies \varphi(T^k(x)) > M \text{ for } k \leq n \implies S_n(x) \geq nM$$

we conclude that

$$\int e^{tS_n(\varphi)} d\mu \geq \mu(\{x : d(x,p) < \frac{r}{\lambda^n}\}) e^{tnM} \geq \frac{cr^d}{\lambda^{dn}} e^{tnM}$$

and therefore

$$\liminf_{n\to\infty} \frac{1}{n} \log \left(\int e^{tS_n(\varphi)} d\mu \right) \geq tM - d \log \lambda \geq L - d \log \lambda$$

Since L was arbitrary, we obtain (3.1). □

Example 3.3. *(i) Suppose $T : X \to X$ is a C^1 map of a compact Riemannian metric space X which preserves a measure μ equivalent to volume and T is exponentially mixing for Hölder functions in the sense that for all φ, ψ which are α-Hölder on X there exist constants C, $0 < \theta < 1$ such that for all $n \geq 0$*

$$|\int \varphi\psi \circ T^n d\mu - \int \varphi d\mu \int \psi d\mu| \leq C\theta^n \|\varphi\|_\alpha \|\psi\|_\alpha$$

Theorem 5.1 of [HNPV13] shows that if B_i is a nested sequence of balls with center a periodic point $p \in X$ and there exists a constant $C_3 > 0$ such that $\mu(B_i) \geq C_3/i$ for all $i > 0$ then μ a.e. $x \in X$ satisfies $T^n x \in B_n$ infinitely often. Thus in this setting if φ is a continuous observable on $X/\{p\}$ such that $\lim_{x\to p} \varphi(p) = \infty$, $\int \varphi d\mu = 0$ and $\varphi > -\rho$ for some $\rho > 0$ then $\{\varphi \circ T^j\}$ does not satisfy exponential large deviations with a rate function.

(ii) In particular, the tent map T of the unit interval $\Omega = [0, 1]$:

$$T(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq \frac{1}{2}; \\ 2x - 1 & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

preserves Lebesgue measure and has exponential decay of correlations for Hölder observables on the system. The function $\varphi(x) = \log(1 - \log x) - \int \log(1 - \log x) dx$ is continuous except at the fixed point 0, and satisfies the assumptions of Theorem 3.1 so does not have exponential large deviations with a rate function.

3.1 Exponential decay of autocorrelations and in L^p, but strictly stretched exponential large deviations.

In this section we consider piecewise expanding C^2 maps of the interval or, more generally, Rychlik maps. For the definition of a Rychlik map (piecewise expanding with countable many branches and additional properties) see [Ryc83] or [BG97, Section 5.4]. In Remark 3.5 we discuss further applications. We consider observables of the form $\varphi(x) = (-\log|x-p|)^\alpha$ where p is periodic. We will show that $\varphi(x)$ has strictly stretched exponential large deviations, despite having exponential decay of autocorrelations and exponential decay in L^p.

Recall that for a μ-preserving map T which is non-singular, the transfer operator $P : L^1(\mu) \to L^1(\mu)$ is defined uniquely by the condition $\int (Pf)g\, d\mu = \int f(g \circ T)\, d\mu$ for all $f \in L^1(\mu)$, $g \in L^\infty(\mu)$.

Theorem 3.4. *Let (T, Ω, μ) be a topologically mixing Rychlik map of the unit interval $\Omega = [0,1]$ and $\varphi(x) := (-\log|x-p|)^\alpha$ where p is periodic for T and $\alpha > 0$. Assume that T is C^1 on a neighborhood of the orbit of p. Then:*

(a) There are $c, C, r > 0$ such that for all $\varepsilon > 0$ close to zero and $\delta > 0$

$$c\ \exp(-rn^{\frac{1}{1+\alpha}}) \le \mu\left(S_n(\varphi) - n\int \varphi d\mu \ge n\varepsilon\right) \le C\exp(-n^{\frac{1}{1+\alpha}-\delta}) \qquad \text{as } n \to \infty$$

Therefore, for any $\varepsilon > 0$ close to zero,

$$\lim_{n\to\infty} \frac{\log\left[-\log\mu\left(S_n(\varphi) - n\int\varphi d\mu \ge n\varepsilon\right)\right]}{\log n} = \frac{1}{1+\alpha}$$

(b) φ has exponential decay of autocorrelations: there exists $0 < \theta < 1$ and $C > 0$ such that

$$\left|\int (\varphi\circ T^n - \overline{\varphi})(\varphi - \overline{\varphi})\, d\mu\right| \le Ce^{-\theta n} \qquad \text{for all } n.$$

See Proposition 3.10 for a more general setting.

(c) $\|P^n(\varphi - \overline{\varphi})\|_{L^p} \to 0$ exponentially fast for each $p \in [1,\infty)$; see (a) of Proposition 3.13 for details and more results.

Remark 3.5. *By equation (3.2), for p a fixed point it suffices in (a) to take $r > (\int\varphi d\mu + \varepsilon)^{1/\alpha} + \log|T'(p)|$.*

More generally than Rychlik maps, one can consider mixing AFN maps introduced in [Zwe98]. For more about these, see [KS19, Section 1.3]. These maps satisfy Property $\mathfrak{D}$ introduced by Schindler [Sch15] and described also in [KS19], see the discussion preceding Proposition 6.1 in the Appendix.

The mixing assumption is only used for the upper bound in (a). For the lower bound in (a) it suffices that μ be an a.c.i.p. whose density is bounded below in a neighborhood of the orbit of p.

Remark 3.6. *For example consider $\varphi(x) = -\log x$, an observable defined on the tent map $(T, [0,1], \text{Lebesgue})$. Clearly φ is integrable, $\int\varphi dx = 1$, φ has moments of all orders and $\mathbb{E}[e^{t\varphi}] = \int e^{tx}e^{-x}dx$ exists for $|t| < 1$.*

This satisfies the assumptions of Theorem 3.1, see Example 3.3 of the previous section, and hence does not satisfy exponential large deviations with a rate function.

However, if X_i is a sequence of i.i.d. random variables with the same distribution function as φ and $S_n = \sum_{j=1}^{n} X_j$, then (exponential) large deviations with a rate function holds: for $0 < \varepsilon < 1$

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{S_n}{n} > 1+\varepsilon\right) = -\varepsilon + \log(1+\varepsilon)$$

This follows from [Dur10, Theorem 2.6.3]; e.g., note that φ has an exponential distribution as $\mu(\varphi > t) = e^{-t}$ and see [Dur10, Example 2.6.2].

Remark 3.7. *By [AFLV11, Proposition 4.1, part (2)] applied to the tent or doubling map on the unit interval (so their θ is 1), there exist constants C_1, C_2 such that for $\varphi(x) := -\log|x-z|$, $\{\varphi \circ T^j\}$ has stretched exponential deviations at rate at least as fast as $C_1 e^{-C_2 n^{\frac{1}{9}-\delta}}$ for any $\delta > 0$.*

However, for z a fixed point of the map, Theorem 3.4 above with $\alpha = 1$ gives an upper bound $C\exp(-n^{\frac{1}{2}-\delta})$ and a lower bound $c\exp(-[(\int \varphi d\mu + \varepsilon) + \log 2 + \delta]n^{\frac{1}{2}})$, $\delta > 0$.

Remark 3.8. *In the setting of non-uniformly expanding maps Araújo [Ara07a, Theorem 3.1] gives an asymptotic condition on observables of form $\varphi(x) = -\log|x-p|$, called exponential slow recurrence to p, which implies exponential large deviations for φ. He also proves that singular points for certain Lorenz-like maps with unbounded derivatives satisfy this condition, which is in general difficult to check.*

Proof of Theorem 3.4. We will only prove (a) here, for parts (b) and (c) see Proposition 3.10 and Proposition 3.13.

We obtain the lower bound from an easy direct computation. For the upper bound we use recent results of Tanja Schindler [Sch15, Lemma 6.15], see Proposition 6.1 in the Appendix. Related estimates are given in [KS19].

We mention here that with the argument of Schindler it does not seem possible to obtain exponential decay of correlations by assuming observables have slower growth rate than $-\log^+(x)$.

Beyond Property $\mathfrak{D}$, we use the fact that a mixing Rychlik map has an invariant density in $BV[(0,1])$, bounded away from zero.

Without loss of generality we will take p to be a fixed point.

Denote $S_n(\varphi)$ by S_n for short. Denote the absolutely continuous invariant probability of T by $d\mu = h\,dx$, with $0 < m \le h \le M$.

Lower bound: We will discuss only the case $T'(p) > 0$ and $p \neq 1$, the others being similar. Pick $\lambda > 1$ such that $T'(x) \le \lambda$ in a neighborhood of p.

Consider $x \in [p, p + e^{-rn^\omega}]$ with $r, \omega > 0$ to be determined later. Then $T^j(x) \in [p, p + \lambda^j e^{-rn^\omega}]$ so $\varphi(T^j(x)) \ge [rn^\omega - j\log(\lambda)]^{1/\alpha}$ as long as all the iterates stay close to p, and therefore, for $K \le n$,

$$S_n(x) \ge \sum_{j=1}^{K} \varphi(T^j(x)) \ge K[rn^\omega - K\log(\lambda)]^{1/\alpha} = \widetilde{K}$$

Thus

$$\mu(S_n - n\overline{\varphi} \geq n\varepsilon) \geq \mu([p, p + e^{-rn^\omega}]) \geq me^{-rn^\omega}$$

as long as

$$\widetilde{K} \geq n(\overline{\varphi} + \varepsilon) \iff rn^\omega \geq \left(\frac{n(\overline{\varphi} + \varepsilon)}{K}\right)^{1/\alpha} + K \log \lambda \qquad \text{and} \qquad \lambda^K e^{-rn^\omega} \ll 1$$

The choice

$$\omega = \frac{1}{1+\alpha}, \qquad K = n^\omega, \qquad r > (\overline{\varphi} + \varepsilon)^{1/\alpha} + \log \lambda \tag{3.2}$$

satisfies both conditions as $n \to \infty$.

Upper bound: We define a sequence of functions (truncated φ) by

$$g_n(x) = \begin{cases} 0 & \text{if } d(x,p) \leq e^{-n^\beta}; \\ \varphi(x) & \text{if } d(x,p) > e^{-n^\beta}. \end{cases}$$

Then $\|g_n\|_{BV} \leq 3|g_n|_\infty = 3n^{\beta\alpha}$. We calculate $\mathbb{E}(g_n) = \mathbb{E}(\varphi) + O(e^{-n\beta}(n\beta)^\alpha)$. Hence, by Schindler's estimate [Sch15] (see Proposition 6.1 in the Appendix), if $G_n := \sum_{j=1}^n g_n \circ T^j$ and $E_n := \int G_n d\mu$ then for any sequence ζ_n tending to zero (we take $\zeta_n = (\log \log n)^{-1}$), for n sufficiently large

$$\mu(|G_n - E_n| > \varepsilon E_n) \leq 2 \exp\left(-n \frac{\mathbb{E}(g_n)}{n^{\beta\alpha}} \zeta_n\right)$$
$$= 2 \exp\left(-n^{1-\beta\alpha} \frac{\overline{\varphi} + O(e^{-n\beta}(n\beta)^\alpha)}{\log \log n}\right) \leq 2e^{-n^{1-\beta\alpha-\delta}}$$

for $\delta > 0$ as $n \to \infty$. Taking into account that $\mu(S_n \neq G_n) \leq O(ne^{-n^\beta})$, we choose β to maximize $\min(1 - \beta\alpha, \beta)$; this gives $\beta = \frac{1}{1+\alpha}$ and thus the claimed upper bound. $\square$

We establish next exponential decay of auto correlations.

Remark 3.9. *Since for the above maps T the transfer operator P has a spectral gap on BV, it follows easily that for any $z \in [0,1]$ the function $\varphi(x) := (-\log d(x,z))^\alpha$, $\alpha > 0$, has exponential decay against BV-observables.*

Proposition 3.10. *Assume $T : [0,1] \to [0,1]$ has an a.c.i.p. μ with density bounded above, and its transfer operator P has a spectral gap on BV. Then $\varphi(x) := (-\log d(x,z))^\alpha$, $\alpha > 0$, has exponential decay of autocorrelations.*

Namely, let r be the spectral radius of P on $BV \cap \{f \mid \int f d\mu = 0\}$. Then, for every $\beta < \min\{-\log r, \frac{1}{2}\}$ there exists a constant $C_\beta > 0$ such that

$$\left| \int (\varphi \circ T^n - \overline{\varphi})(\varphi - \overline{\varphi})\, d\mu \right| \leq C_\beta e^{-\beta n} \qquad \textit{for all } n.$$

Proof. Define a sequence of truncations

$$h_n(x) := \min\{M_n, \varphi(x)\}, \qquad \delta_n := \varphi - h_n, \qquad \overline{\varphi} := \int \varphi d\mu, \qquad \overline{h_n} := \int h_n d\mu$$

with M_n to be determined later. Note that $\varphi \in L^2(\mu)$ since μ has a density bounded above.

Obviously $\|h_n\|_{BV} \leq 3|h_n|_\infty = 3M_n$. Using the spectral gap of the transfer operator P in BV, for every $r < \rho < 1$ there is a $C_P > 0$ such that $|P^n(h)|_\infty \leq \|P^n(h)\|_{BV} \leq C_P\rho^n\|h\|_{BV}$ provided $\int h d\mu = 0$.

Then, using the spectral gap of P on BV, Hölder's inequality and that $|\overline{h_n} - \overline{\varphi}| = |\int \delta_n d\mu| \leq \|\delta_n\|_{L^2}$, and denoting by C all constants that do not change with M_n and n:

$$\begin{aligned}\left|\int \varphi \circ T^n \cdot (\varphi - \overline{\varphi}) d\mu\right| &= \left|\int \varphi \circ T^n \cdot [(h_n - \overline{h_n}) + (\delta_n + \overline{h_n} - \overline{\varphi}) d\mu\right| \\ &\leq \left|\int \varphi \cdot P^n(h_n - \overline{h_n}) d\mu\right| + \|\varphi\|_{L^2}(\|\delta_n\|_{L^2} + |\overline{h_n} - \overline{\varphi}|) \\ &\leq C\rho^n\|\varphi\|_{L^1}\|h_n - \overline{h_n}\|_{BV} + C\|\varphi\|_{L^2}\|\delta_n\|_{L^2}\end{aligned}$$

Since

$$\delta_n(x) \neq 0 \iff d(x,z) \leq M_n^{1/\alpha}$$

and μ has a density bounded above, for any $\delta > 0$ we obtain

$$\|\delta_n\|_{L^2}^2 \leq C \int_{|x-z| \leq M_n^{1/\alpha}} |\log(|x-p|)|^{2\alpha} dx \leq C_\delta e^{-(1-\delta)M_n^{1/\alpha}}$$

and therefore

$$\left|\int \varphi \circ T^n \cdot (\varphi - \overline{\varphi}) d\mu\right| \leq C\|\varphi\|_{L^1} M_n \rho^n + CC_\delta \|\varphi\|_{L^2} e^{-\frac{1}{2}(1-\delta)M_n^{1/\alpha}}$$

Choose now $M_n = n^\alpha$. Hence for every $\beta < \min\{-\log\rho, \frac{1}{2}\}$ there exists a constant $C_\beta > 0$ such that for all n

$$\left|\int (\varphi \circ T^n - \overline{\varphi})(\varphi - \overline{\varphi})\, d\mu\right| \leq C_\beta e^{-\beta n}.$$

□

3.2 Stretched exponential large deviations from exponential decay on Lipschitz functions

In this section we slightly improve estimates of [AFLV11][Proposition 4.1 (b)] to obtain a better stretched exponent decay rate.

Theorem 3.11. *Let (T, X, μ) with $X = [0,1]$ be a dynamical system where μ is a T-invariant a.c.i.p having density bounded above, $d\,\mu/d$ Lebesgue $\leq C_\mu$.*

Consider the observation $\varphi : [0,1] \to \mathbb{R}$, $\varphi(x) := |\log d(x,z)|$ for some $z \in X$.

Assume P, the transfer operator w.r.t. μ, has exponential decay in the space of Lipschitz functions: there are constants $\theta, C_P > 0$ such that

$$\|P^n f\|_{\text{Lip}} \leq C_P e^{-\theta n}\, \|f\|_{\text{Lip}} \;\; \textit{provided} \int f d\mu = 0.$$

Fix $\alpha < 1/5$. Then for $\varepsilon > 0$ close to zero there are $C = C_{\varepsilon,\alpha} > 0$ and $r = r_{\varepsilon,\alpha} > 0$ such that

$$\mu\left(\left|S_n(\varphi - \int \varphi d\mu)\right| > n\varepsilon\right) \leq C\exp(-rn^\alpha)$$

Remark 3.12.

1. *The doubling map satisfies the hypothesis of Theorem 3.11. Therefore, by Theorem 3.4 cannot have decay rate faster than* $\exp(-n^{1/2})$.

2. *In this setting a similar stretched exponential large deviations was also obtained in [AFLV11][Proposition 4.1 (b)], but with rate* $\exp(-n^{1/9})$.

Proof. For simplicity, we take z to be zero, so $\varphi = |\log x|$.

We denote by

$$\widetilde{f} := f - \int f d\mu$$

and when this correction $\int f d\mu$ is $O(1)$ we might ignore it in the estimates.

We introduce three parameters, M_n, C_n and α, whose optimal value will be determined at the end.

Denote $h^{(n)} := \min\{\varphi, M_n\}$. Define $w^{(n)} := \sum_{k\geq 1} P^k \widetilde{h^{(n)}}$ and write $h^{(n)} := g^{(n)} + w^{(n)} \circ T - w^{(n)}$. Since $Pg^{(n)} = 0$, its Birkhoff sums form a martingale.

Recall the Azuma-Hoeffding inequality: if $S_n := \sum_{k=1}^n X_n$ is a martingale whose increments X_n satisfy $|X_n| \leq M_n$ then

$$P(S_n \geq A) \leq \exp\left(-\frac{A^2}{2\sum_{i=1}^n M_i^2}\right)$$

Using Azuma-Hoeffding,

$$\begin{aligned}
\mu(S_n(\widetilde{h^{(n)}}) > 3n\varepsilon) &\leq \mu(S_n(\widetilde{g^{(n)}}) > n\varepsilon) + 2\mu(w^{(n)} > n\varepsilon) \\
&\leq \exp(-\frac{n^2\varepsilon^2}{2n|\widetilde{g^{(n)}}|^2_{L^\infty}}) + 2\mu(w^{(n)} > n\varepsilon)
\end{aligned}$$

and therefore

$$\begin{aligned}
\mu(S_n(\widetilde{\varphi}) > 3n\varepsilon) &\leq \mu(S_n(\widetilde{h^{(n)}}) > 3n\varepsilon) + n\mu(h^{(n)} \neq \varphi) \\
&\leq \exp(-\frac{n^2\varepsilon^2}{2n|\widetilde{g^{(n)}}|^2_{L^\infty}}) + 2\mu(w^{(n)} > n\varepsilon) + C_\mu n e^{-M_n} \qquad (3.3)
\end{aligned}$$

Compute

$$|\widetilde{g^{(n)}}|_{L^\infty} \leq |h^{(n)}|_{L^\infty} + 2|w^{(n)}|_{L^\infty} \qquad (3.4)$$

and, using the exponential decay of P and that $|Pf|_{L^\infty} \leq |f|_{L^\infty}$,

$$\begin{aligned}
|w^{(n)}|_{L^\infty} &= |\sum_{k=1}^\infty P^k \widetilde{h^{(n)}}|_{L^\infty} \\
&\leq \sum_{k=1}^{C_n} |P^k \widetilde{h^{(n)}}|_{L^\infty} + \sum_{k=C_n+1}^\infty \left\| P^k \widetilde{h^{(n)}} \right\|_{\mathrm{Lip}} \\
&\leq C_n M_n + C_P \sum_{k=C_n+1}^\infty e^{-\theta k} \left\| \widetilde{h^{(n)}} \right\|_{\mathrm{Lip}} \\
&\leq C_n M_n + C' e^{-\theta C_n} \left\| \widetilde{h^{(n)}} \right\|_{\mathrm{Lip}} \\
&\leq C_n M_n + C' e^{-\theta C_n} e^{M_n} \qquad (3.5)
\end{aligned}$$

We want to bound each term in (3.3) by $\exp(-n^\alpha)$. For the first term this requires $|\widetilde{g^{(n)}}|_{L^\infty} \approx n^{(1-\alpha)/2}$, hence, in view of (3.4) and (3.5), need

$$C_n M_n \approx n^{(1-\alpha)/2} \qquad -\theta C_n + M_n \lesssim \frac{1-\alpha}{2}\log n \tag{3.6}$$

which lead to

$$M_n = n^{(1-\alpha)/4} \qquad C_n = \frac{1}{\theta} M_n = \frac{1}{\theta} n^{(1-\alpha)/4}$$

Then the second term in (3.3) is zero because $|w^{(n)}|_{L^\infty} \lesssim n^{(1-\alpha)/2}$, and the third term becomes $ne^{-M_n} \approx \exp(-n^{(1-\alpha)/4})$.

Thus the best α comes from $\max\min\{n^\alpha, n^{(1-\alpha)/4}\} = n^{1/5}$ for $\alpha = 1/5$. $\square$

3.3 Exponential decay in L^p for all $p \geq 1$ does not imply exponential large deviations

We now show that exponential decay of the transfer operator in each L^p does not imply exponential large deviations.

The next result is stated for the doubling map, but the proof applies, with the appropriate changes, to any map having an a.c.i.p. with density bounded above, and whose transfer operator has a spectral gap, e.g. on BV.

Proposition 3.13. *Let (T, X, μ) be the doubling map on $X = [0, 1]$, P its transfer operator with respect to the Lebesgue measure. Denote by $e^{-\theta}$ the exponential decay rate of its transfer operator on BV functions of mean zero, and by C_T the constant that is involved (see (3.7) for the precise meaning of θ and C_T).*

(a) Let $\varphi_1(x) := (-\log x)^\alpha$, $\alpha > 0$. Then for all $p \geq 1$, there is a constant $C_{\alpha p}$ such that

$$\left\| P^n \left(\varphi_1 - \int \varphi_1 dx \right) \right\|_{L^p} \leq e^{-\theta n}(C_T M_n + C_{\alpha p} M_n^\alpha), \text{ where } M_n = np\theta.$$

In particular, the decay is exponential with rate $e^{-\beta}$ for any $\beta < \theta$. The constant $C_{\alpha p}$ comes from (3.8).

(b) Let $\varphi_2(x) := x^{-\alpha}$ with $0 < \alpha < 1$. Then for $1 \leq p < 1/\alpha$

$$\left\| P^n \left(\varphi_2 - \int \varphi_2 dx \right) \right\|_{L^p} \lesssim e^{-n\theta(1-\alpha p)} \text{ as } n \to \infty$$

Proof. The transfer operator P is a contraction in each L^p, $p \geq 1$, and has exponential decay on functions in BV with mean zero; so can find $C_T, \theta > 0$ such that for $f \in L^p$ [1]

$$\|Pf\|_p \leq \|f\|_p, \qquad \left\| P^n \left(\varphi - \int \varphi dx \right) \right\|_{BV} \leq C_T e^{-\theta n} \left\| P^n \left(\varphi - \int \varphi dx \right) \right\|_{BV} \tag{3.7}$$

[1] These inequalities also hold when f is not in L^p or BV, in that case the RHS is infinity.

For φ_k, $k = 1, 2$, substract for simplicity its mean, and denote $\widetilde{\psi} := \varphi_k - m_k$ where $m_k = \int \varphi_k dx$, and let $\psi_n := \min\{\widetilde{\psi}, M_n\}$; in what follows we will ignore this constant m_k, it does not change much.

Then $\|\psi_n\|_{BV} = M_n$ (more precisely, $M_n + m_k$), and

$$\begin{aligned} \|P^n\widetilde{\psi}\|_p &\leq \|P^n\psi_n\|_p + \|P^n(\widetilde{\psi} - \psi_n)\|_p \\ &\leq \|P^n\psi_n\|_{BV} + \|\widetilde{\psi} - \psi_n\|_p \leq C_T e^{-\theta n}\|\psi_n\|_{BV} + \|\widetilde{\psi} - \psi_n\|_p \\ &\leq C_T e^{-\theta n} M_n + \|\widetilde{\psi} - \psi_n\|_p \end{aligned}$$

Now compute that for $\varphi_1(x) := (-\log x)^\alpha$, $\alpha > 0$ (ignoring again m_1):

$$\begin{aligned} \|\widetilde{\psi} - \psi_n\|_p &= \left(\int_{\varphi_1 > M_n} |\varphi_1 - M_n|^p dx\right)^{1/p} \leq \left(\int_{0<x<e^{-M_n}} |\log(x)|^{\alpha p} dx\right)^{1/p} \\ &\leq C_{\alpha p}\left(x|\log(x)|^{\alpha p}\Big|_{x=0}^{x=e^{-M_n}} dx\right)^{1/p} = C_{\alpha p} e^{-M_n/p} M_n^\alpha \end{aligned} \tag{3.8}$$

Take $M_n = np\theta$ to get the desired decay in L^p.

Doing the same for $\varphi_2(x) := x^{-\alpha}$, $0 < \alpha < 1$, $1 \leq p < 1/\alpha$:

$$\|\widetilde{\psi} - \psi_n\|_p \leq \left(\int_{0<x<M_n^{-1/\alpha}} x^{-\alpha p} dx\right)^{1/p} = \frac{1}{(1-\alpha p)^{1/p}} M_n^{1-\frac{1}{\alpha p}}$$

Take $M_n = \exp(p\alpha\theta n)$ to get exponential decay at rate $e^{-\theta(1-\alpha p)}$. □

4 Exponential large deviations without a rate function for bounded observables.

Examples exist in the literature [Bra89, OP88, BS93, Chu11] of stationary processes which have exponential large deviations but a rate function does not exist. In particular there is an example of a mean zero bounded function f taking only 3 values on an aperiodic recurrent Markov chain (X_n) with a countable state space such that the system has exponential large deviations but does not have a rate function. In this example defining $S_n = \sum_{j=0}^{n-1} f(X_j)$ for all $\varepsilon > 0$, there exist constants $C(\varepsilon)$, $0 < \gamma < 1$ such that $\mathbb{P}(|\frac{S_n}{n}| > \varepsilon) \leq C(\varepsilon)e^{-\gamma n}$, giving exponential convergence in the strong law of large numbers yet there is no rate function controlling the rate of decay. We show in the next section that in these examples f is a coboundary and there exists $\psi \in L^2(\mathbb{P})$ such that $f(X_j) = \psi(X_{j+1}) - \psi(X_j)$ for all $j \geq 0$ so that the example is degenerate (the variance $\sigma^2 = 0$). The assumption that the observable is not a coboundary (and hence that $\sigma^2 > 0$) is made in the statements of the theorems establishing large deviations with rate functions in [MN08].

According to Bryc and Smolenski [BS93] the idea of the example was due to Bradley [Bra89] and also adapted by Orey and Pelikan [OP88, Example 4.1]. Bradley [Bra89] produced an example of a stationary, pairwise independent, absolutely regular stochastic process for which the central limit theorem does not hold. Orey and

Pelikan presented this system as an example of a strongly mixing shift for which the large deviation principle with rate function fails. Bryc and Smolenski showed that there is in fact also an exponential convergence in the strong law of large numbers. Bryc and Smolenski's work was recast by Chung [Chu11] into dynamical systems language, and the system was expressed as a Young Tower (F, Δ, ν). We will also recast as a dynamical system and show that f is a coboundary, in fact $f = \psi \circ F - \psi$ where ψ is unbounded but $\psi \in L^2$. This seems to have been overlooked in the literature. In fact if f were not a coboundary the example would contradict results of [MN08, RBY08] which imply that any bounded Lipschitz function on a Young Tower with exponential tails which is not a coboundary has exponential large deviations with a rate function. In Section 4.1 we also present the example in a dynamical setting following Bryc and Smolenski's notation and overall presentation and give an explicit coboundary for f. As far as we know there is no example of a bounded observable on a dynamical system, with non-zero variance, which has exponential large deviations and yet no rate function.

4.1 The example of Bryc and Smolenski

Let Δ_0 (which for concreteness we will identify with the unit interval $[0,1]$) be the base of a Young Tower Δ with Δ_0 partitioned into intervals $\Lambda_0, \Lambda_1, \ldots, \Lambda_k \ldots$,. Let $n(0) = 1/2$, $n(k) = 12^k$, $p_k = Ce^{-\frac{n(k)}{2}}$ where $C^{-1} = \sum_{k=0}^{\infty} e^{-\frac{n(k)}{2}}$ is a normalization constant to ensure the Tower has a probability measure. Each Λ_k will have Lebesgue measure p_k and we define the Tower return time function on Λ_k as $R_{\Lambda_k} := R(k) = 2n(k)$. We now build the Tower Δ above the base. We write $\Lambda_{k,0} := \Lambda_k$ and define, for $0 \leq j \leq R(k) - 1$ the levels $\Lambda_{k,j}$ of the Tower lying above Λ_k by

$$\Delta = \bigcup_{k \in N^+, 0 \leq j \leq R_k - 1} \{(x, j) : x \in \Lambda_{0,k}\}$$

with the tower map $F : \Delta \to \Delta$ given by

$$F(x, j) = \begin{cases} (x, j+1) & \text{if } x \in \Lambda_{k,0}, j < R(k) - 1 \\ (T_k x, 0) & \text{if } x \in \Lambda_{k,0}, j = R(k) - 1 \end{cases}.$$

where T_k has constant derivative and maps the interval $\Lambda_{k,0}$ bijectively onto Δ_0. We define F on $\Lambda_{0,0}$ by requiring that F map $\Lambda_{0,0}$ bijectively onto Δ_0. Note that $m(\Lambda_{0,0}) = Ce^{-1/4}$. This requirement on the height $R(0)$ of $\Lambda_{0,0}$ to be 1 is to ensure aperiodicity.

We lift Lebesgue measure m from the base to the Tower to obtain a measure $\tilde{\nu}$ on Δ and then define a probability measure $\nu = \tilde{\nu}$ by normalization. The map F preserves ν and is exponentially mixing for a Banach space of observables on Δ [You98].

If $k \neq 0$ we define $f : \lambda_{k,j} \to \{-1, 0, 1\}$ by

$$f(x, j) = \begin{cases} 1 & \text{if } x \in \Lambda_k, j \leq n(k) - 1 \\ -1 & \text{if } x \in \Lambda_k, n(k) \leq j \leq 2n(k) - 1 \end{cases}.$$

if $k = 0$ we take $f(0, 0) = 0$. This is the example model of [Bra89, OP88, BS93, Chu11].

Now define a function ψ, which will be a coboundary for f, by

$$\psi(x,j) = \begin{cases} j & \text{if } x \in \Lambda_k, 0 \le j \le n(k) \\ 2n(k) - j & \text{if } x \in \Lambda_k, n(k) < j \le 2n(k) - 1 \end{cases}.$$

and we take $\psi(0,0) = 0$.

It is easy to check that

$$f = \psi \circ F - \psi$$

Hence $S_n(f) = \psi \circ F^n - \psi$. We remark that, as $f = \psi \circ F - \psi$ is a coboundary and $\psi \in L^2$, $S_n(f)$ has zero-variance and does not satisfy a non-trivial central limit theorem.

Since

$$\nu(\psi \circ F^n > n) = \nu(\psi > n) \sim e^{-n/2}$$

it is easy to see that $\nu(|S_n/n| > \varepsilon)$ decays exponentially for any ε (and obtain explicit estimates).

In [Var84, Theorem 2.2] Varadhan shows that large deviations with rate function fails if $\frac{1}{n} \log \mathbb{E}(\exp(S_n))$ does not have a limit as $n \to \infty$ and it is easy to show (see also [BS93]) that $\frac{1}{n} \log \mathbb{E}(e^{S_n})$ has no limit in this example.

For example consider points in $x \in \Lambda_{k,0}$ where $\nu(\Lambda_{k,0}) = Ce^{-\frac{n(k)}{2}}$. Then $S_{n(k)}(x) = \psi \circ F^{n(k)} = n(k)$ so that $e^{S_{n(k)}(x)} = e^{n(k)}$. Hence

$$\frac{1}{n(k)} \log \mathbb{E}(e^{S_{n(k)}})) = \frac{1}{2n(k)} \log(Ce^{-\frac{n(k)}{2}} e^{n(k)}) = \frac{1}{4} + \frac{\log C}{n(k)}.$$

Thus $\limsup \frac{1}{n} \log \mathbb{E}(e^{S_n}) \ge \frac{1}{4}$. Similar considerations show that $\liminf \frac{1}{n} \log \mathbb{E}(e^{S_n}) \le 0$ and consequently $\frac{1}{n} \log \mathbb{E}(e^{S_n})$ has no limit.

5 Discussion and open problems.

Large deviations theory is well developed in the probabilistic setting and there are many standard texts [DZ10, Ell06, Var84]. Fundamental work was done by Donsker-Varadhan [DV75, DV76, DV83] for certain random processes, including Markov chains and Brownian motion.

In the discrete case, assume given an identically distributed process $X_0, X_1, \dots$ taking values in a Polish space $\mathcal{X}$, and φ a suitable observable on $\mathcal{X}$. A level I Large Deviation Principle estimates the (exponentially decaying with the sample size) probability of sample averages being away from the mean. The stronger level II Large Deviation Principle estimates the probability of the empirical distribution being in a given set of probability distributions on $\mathcal{X}$ that does not contain the common distribution of the X_k's. See e.g. [Var08].

In the case of a dynamical system we describe the case of a level I large deviation with rate function; the sample mean is replaced by Birkhoff averages, namely

$$\lim_{n\to\infty} \frac{1}{n} \log \mu(S_n(\varphi) \ge n\alpha) = -I(\alpha) \tag{5.1}$$

for α close to the mean $\int \varphi d\mu$ of φ, where $I(\alpha)$ is strictly convex and vanishes only at $\int \varphi d\mu$. Such a function is called a *rate function* and is often characterized in terms of thermodynamic quantities.

In the setting of dynamical systems with some degree of hyperbolicity there have been many results on large deviations going back to the 1980s. Takahashi [Tak84] has a very first attempt to study large deviation in dynamics. Orey [Ore86] proves large deviation results for Markov chains, and applies it to one-dimensional dynamical systems; see also Denker [Den92]. Orey and Pelikan [OP89] obtained a large deviation theorem for Hölder observables on Anosov diffeomorphisms. Lopes [Lop90] using similar ideas obtained large deviations with respect to the measure of maximal entropy in the setting of an expanding rational map of the Riemann sphere of degree $d \geq 2$ restricted to its Julia set. Kifer [Kif90] proved large deviations results for uniformly hyperbolic systems, including flows. Young [You90] gave general theorems for continuous transformations on a compact metric space and more detailed quantitative estimates in the setting of SRB attractors, Axiom A attractors and subshifts of finite type. Takahashi [Tak87] has similar results to Young, using techniques from thermodynamical formalism and at about the same time. Kifer [Kif90] and Young [You90] formulated quite general large deviation principles for dynamical systems; for example Kifer obtained the upper bound half of (5.1) for uniformly partially hyperbolic dynamical systems. However, these results yield strong conclusions (in particular (5.1)) only if it is known that there is a unique equilibrium measure for the underlying map. There is also work of Waddington [Wad96] on large deviations for continuous observables on Anosov flows.

Thus large deviations theory for uniformly hyperbolic (Axiom A) dynamical systems, for both discrete and continuous time, is fairly well-understood by these results. Furthermore, when X is an Axiom A attractor and μ is an SRB measure, then μ can be replaced by Lebesgue measure in (5.1).

Following these results there were also level II large deviation principles by Grigull [Gri93] (see also the survey by Denker [Den96]) for Hölder observables on parabolic rational maps. For intermittent type maps Pollicott, Sharp and Yuri [PSY98] have upper bounds for sums of the preimages weighted by the derivative. For a general class of one-dimensional unimodal maps, Keller and Nowicki [KN92] obtained large deviations results (5.1) for observables of bounded variation in terms of Lebesgue measure.

In the late 1990s the study of the statistical properties of non-uniformly hyperbolic dynamical systems was advanced significantly by L-S. Young [You98, You99] with the application of an inducing scheme formulated as a Young Tower. Using the structure of a Young Tower and transfer operator techniques Melbourne and Nicol [MN08] obtained large deviation estimates for a large class of nonuniformly hyperbolic systems: namely those modelled by Young towers with summable decay of correlations. In the case of exponential decay of correlations, they obtained exponential large deviation estimates given by a rate function. In the case of polynomial decay of correlations, using martingale techniques, they obtained polynomial large deviation estimates, and exhibited examples where these estimates are essentially optimal. For Hölder observables they obtained exponential estimates in situations where the space of equilibrium measures is not known to be a singleton, as well as polynomial estimates in situations where there is not a unique equilibrium measure. Melbourne improved the range of parameters for which the estimates hold in [Mel09] and also gave moderate deviations results. Roughly at the same time Rey-Bellet and Young [RBY08] proved large deviation and moderate deviation results for dynamical systems modeled by Young towers with exponential tails. The proofs of [MN08] and [RBY08] are largely based on quasicompactness of the transfer operator on an appropriately chosen Banach space. The manuscript of Hennion and Hervé [HH01] gives a good account of this theory. In the setting of intermittent maps Pollicott and Sharp [PS09] established level II large deviations for Hölder observables and among other results showed that if $\varphi(0) > \varepsilon$ (here 0 is the indifferent fixed point) then an exponential large deviation holds.

Other results in the non-uniformly hyperbolic setting include those by Araújo and Pacifico [AP06] who obtained large deviation results, in terms of Lebesgue measure, for continuous functions over non-uniformly expanding maps with non-flat singularities or criticalities and for certain partially hyperbolic non-uniformly expanding attracting sets. Araújo [Ara07b] extended these results to obtain large deviation bounds for continuous functions on suspension semiflows over a non-uniformly expanding base transformation with non-flat singularities or criticalities (including semiflows modeling the geometric Lorenz flow and the Lorenz flow). But these results in [Ara07b, AP06] yield strong conclusions only when there is a unique equilibrium measure. Several recent works have obtained large deviations under weakened notions of specification. A partial list includes Bomfim and Varandas [BV19], Varandas [Var12], Pfister and Sullivan [PS05].

Aaronson and Denker [AD90, Theorem 3] give exponential rate large deviations for infinite measure preserving transformations.

We note that some approaches, such as those based on specification [You90], allow large deviation results for continuous observables to be established while transfer operator techniques usually require more regularity of the observable. On the other hand the thermodynamical approach of Young [You90] in general requires the uniqueness of equilibrium states and specification to hold.

Open problems.

A theory of large deviations for unbounded observables on hyperbolic dynamical systems has yet to be developed. There are many settings in which such observables arise naturally. For example is there a rate function for $-\log DT_u$, the derivative of the Jacobian in the unstable direction when the underlying system is the billiard map associated to a planar Sinai dispersing billiard? In other situations we may wish for an estimate of the rate of convergence of finite time Lyapunov exponents to their spatial average when the underlying map or flow has an unbounded derivative. For dynamical systems that are polynomially mixing close to optimal large deviations estimates can probably be obtained using martingale methods, so roughly speaking the problem is to obtain optimal large deviations estimates for an unbounded, integrable observable on an exponentially mixing dynamical system. Here are some more concrete questions that arise from this paper.

(1) Does the observable $\varphi(x) = -\log d(x,p)$ on a uniformly expanding map of the interval satisfy exponential large deviations for a full measure set of p? If p is a non-periodic point does the observable $\varphi(x) = -\log d(x,p)$ satisfy exponential large deviations?

(2) Are there conditions on the decay rate of the tail of the distribution $\mu(\varphi > t)$ which imply exponential large deviations if φ is an unbounded observable on a uniformly expanding map? Theorem 3.1 shows that we cannot expect to have exponential large

deviations with a rate function.

(3) Is there an example of a non-degenerate bounded observable φ, $\varphi \neq \psi - \psi \circ T$, on a (smooth) dynamical system which satisfies exponential large deviations but does not have a rate function?

In addition, obtaining type II large deviations in this setting is also an open problem; some results were obtained by Grigull [Gri93] in his dissertation.

6 Appendix

We describe Schindler's result that we use, [Sch15, Lemma 6.15], which states that Property $\mathfrak{D}$ for $(\Omega, \mathcal{B}, T, \mu, \mathcal{F}, \|\cdot\|, \chi)$ implies Property $\mathbf{A}$ for $(\chi \circ T^{n-1})_n$, with Property $\mathfrak{D}$ defined in [Sch15, Definition 1.7 in Section 1.2] and Property $\mathbf{A}$ in [Sch15, Definition 2.2]. We state this result only for the situation we are interested in, namely the Banach space $\mathcal{F}$ being $BV([0,1))$. We will use $\|f\|_{BV} := \mathrm{var}(f) + |f|_\infty$ with var denoting the total variation seminorm.

For a function $\chi : \Omega \to \mathbb{R}$ and $\ell \in \mathbb{R}$, denote its truncation by $\chi^\ell := \chi \cdot \mathbf{1}_{\{\chi \leq \ell\}}$.

Proposition 6.1 ([Sch15, Lemma 6.15]). *Let $T : [0,1] \to [0,1]$ be a transformation and μ a T-invariant probability measure for which T is non-singular; denote by P the transfer operator associated to T with respect to the measure μ. Assume:*

(a) the T-invariant probability measure μ is mixing;

(b) P is bounded on $BV[(0,1)]$ and has a spectral gap on BV (that is, the spectral radius of P restricted to the codimension one subspace $BV \cap \{f \mid \int f d\mu = 0\}$ is less than 1);

(c) $\chi \in L^1(\mu)$ with $\chi \geq 0$ and $|\chi|_{L^\infty} = \infty$;

(d) there exists $C > 0$ such that for all $\ell > 0$,

$$\left\|\chi^\ell\right\|_{BV} \leq C \cdot \ell$$

where $X^f := X \cdot \mathbf{1}_{X \leq f}$.

Then, for every sequence $(\xi_n)_{n\in\mathbb{N}}$ tending monotonically to zero and for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for every positive valued sequence (f_m):

$$n \geq N \implies \mathbb{P}\left(\left|T_n^{f_n} - \mathbb{E}\left(T_n^{f_n}\right)\right| > \varepsilon \mathbb{E}\left(T_n^{f_n}\right)\right) \leq 2\exp\left(-\xi_n \frac{\mathbb{E}\left(T_n^{f_n}\right)}{f_n}\right)$$

where $T_n^{f_n} := \sum_{k=1}^n \chi^{f_n} \circ T^k$ is the Birkhoff sum of n truncated terms.

[Note that if $\chi \geq 0$ is bounded then we get this estimate with $f_n = |\chi|_{L^\infty(\mu)}$.]

References

[AD90] Jon Aaronson and Manfred Denker, *Upper bounds for ergodic sums of infinite measure preserving transformations*, Trans. Amer. Math. Soc. **319** (1990), no. 1, 101–138.

[AFLV11] José F. Alves, Jorge M. Freitas, Stefano Luzzatto, and Sandro Vaienti, *From rates of mixing to recurrence times via large deviations*, Adv. Math. **228** (2011), no. 2, 1203–1236.

[AP06] V. Araújo and M. J. Pacifico, *Large deviations for non-uniformly expanding maps*, J. Stat. Phys. **125** (2006), no. 2, 415–457.

[Ara07a] Vítor Araújo, *Large deviations bound for semiflows over a non-uniformly expanding base*, Bull. Braz. Math. Soc. (N.S.) **38** (2007), no. 3, 335–376.

[Ara07b] Vítor Araújo, *Large deviations bound for semiflows over a non-uniformly expanding base*, Bull. Braz. Math. Soc. (N.S.) **38** (2007), no. 3, 335–376.

[BG97] Abraham Boyarsky and Paweł Góra, *Laws of chaos*, Probability and its Applications, Birkhäuser Boston, Inc., Boston, MA, 1997, Invariant measures and dynamical systems in one dimension.

[Bra89] Richard C. Bradley, *A stationary, pairwise independent, absolutely regular sequence for which the central limit theorem fails*, Probab. Theory Related Fields **81** (1989), no. 1, 1–10.

[BS93] W. Bryc and W. Smoleński, *On the convergence of averages of mixing sequences*, J. Theoret. Probab. **6** (1993), no. 3, 473–483.

[BV19] Thiago Bomfim and Paulo Varandas, *The gluing orbit property, uniform hyperbolicity and large deviations principles for semiflows*, J. Differential Equations **267** (2019), no. 1, 228–266.

[Chu11] Yong Moo Chung, *Large deviations on Markov towers*, Nonlinearity **24** (2011), no. 4, 1229–1252.

[CR79] M. Csörgő and P. Révész, *How big are the increments of a Wiener process?*, Ann. Probab. **7** (1979), no. 4, 731–737.

[Den92] M. Denker, *Large deviations and the pressure function*, Transactions of the Eleventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes., Czech Academy of Sciences, 1992, held in Prague, August 26-31, 1990 (preprint: Mathematica Gottingiensis 38, 1988), pp. 21–23.

[Den96] Manfred Denker, *Probability theory for rational maps*, Probability theory and mathematical statistics (St. Petersburg, 1993), Gordon and Breach, Amsterdam, 1996, pp. 29–40.

[DGM10] J. Dedecker, S. Gouëzel, and F. Merlevède, *Some almost sure results for unbounded functions of intermittent maps and their associated Markov chains*, Ann. Inst. Henri Poincaré Probab. Stat. **46** (2010), no. 3, 796–821.

[DK07] Manfred Denker and Zakhar Kabluchko, *An Erdös-Rényi law for mixing processes*, Probab. Math. Statist. **27** (2007), no. 1, 139–149.

[DN13] Manfred Denker and Matthew Nicol, *Erdös-Rényi laws for dynamical systems*, J. Lond. Math. Soc. (2) **87** (2013), no. 2, 497–508.

[Dur10] Rick Durrett, *Probability: theory and examples*, fourth ed., Cambridge Series in Statistical and Probabilistic Mathematics, vol. 31, Cambridge University Press, Cambridge, 2010.

[DV75] M. D. Donsker and S. R. S. Varadhan, *Asymptotic evaluation of certain Markov process expectations for large time. I. II*, Comm. Pure Appl. Math. **28** (1975), 1–47; ibid. 28 (1975), 279–301.

[DV76] ———, *Asymptotic evaluation of certain Markov process expectations for large time. III*, Comm. Pure Appl. Math. **29** (1976), no. 4, 389–461.

[DV83] ———, *Asymptotic evaluation of certain Markov process expectations for large time. IV*, Comm. Pure Appl. Math. **36** (1983), no. 2, 183–212.

[DZ10] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010, Corrected reprint of the second (1998) edition.

[Ell06] Richard S. Ellis, *Entropy, large deviations, and statistical mechanics*, Classics in Mathematics, Springer-Verlag, Berlin, 2006, Reprint of the 1985 original.

[ER70] Paul Erdős and Alfréd Rényi, *On a new law of large numbers*, J. Analyse Math. **23** (1970), 103–111.

[Gri93] J Grigull, *Große Abweichungen und Fluktuationen für Gleichgewichtsmaße rationaler Abbildungen*, PhD dissertation, Georg-August-Universität zu Göttingen, 1993.

[HH01] Hubert Hennion and Loïc Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Lecture Notes in Mathematics, vol. 1766, Springer-Verlag, Berlin, 2001.

[HNPV13] Nicolai Haydn, Matthew Nicol, Tomas Persson, and Sandro Vaienti, *A note on Borel-Cantelli lemmas for non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **33** (2013), no. 2, 475–498.

[Kif90] Yuri Kifer, *Large deviations in dynamical systems and stochastic processes*, Trans. Amer. Math. Soc. **321** (1990), no. 2, 505–524.

[KN92] Gerhard Keller and Tomasz Nowicki, *Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps*, Comm. Math. Phys. **149** (1992), no. 1, 31–69.

[KS19] Marc Kesseböhmer and Tanja Schindler, *Strong laws of large numbers for intermediately trimmed Birkhoff sums of observables with infinite mean*, Stochastic Process. Appl. **129** (2019), no. 10, 4163–4207.

[Lop90] Artur O. Lopes, *Entropy and large deviation*, Nonlinearity **3** (1990), no. 2, 527–546.

[Mel09] Ian Melbourne, *Large and moderate deviations for slowly mixing dynamical systems*, Proc. Amer. Math. Soc. **137** (2009), no. 5, 1735–1741.

[MN08] Ian Melbourne and Matthew Nicol, *Large deviations for nonuniformly hyperbolic systems*, Trans. Amer. Math. Soc. **360** (2008), no. 12, 6661–6676.

[OP88] Steven Orey and Stephan Pelikan, *Large deviation principles for stationary processes*, Ann. Probab. **16** (1988), no. 4, 1481–1495.

[OP89] ______, *Deviations of trajectory averages and the defect in Pesin's formula for Anosov diffeomorphisms*, Trans. Amer. Math. Soc. **315** (1989), no. 2, 741–753.

[Ore86] Steven Orey, *Large deviations in ergodic theory*, Seminar on stochastic processes, 1984 (Evanston, Ill., 1984), Progr. Probab. Statist., vol. 9, Birkhäuser Boston, Boston, MA, 1986, pp. 195–249.

[PS05] C.-E. Pfister and W. G. Sullivan, *Large deviations estimates for dynamical systems without the specification property. Applications to the β-shifts*, Nonlinearity **18** (2005), no. 1, 237–261.

[PS09] Mark Pollicott and Richard Sharp, *Large deviations for intermittent maps*, Nonlinearity **22** (2009), no. 9, 2079–2092.

[PSY98] Mark Pollicott, Richard Sharp, and Michiko Yuri, *Large deviations for maps with indifferent fixed points*, Nonlinearity **11** (1998), no. 4, 1173–1184.

[RBY08] Luc Rey-Bellet and Lai-Sang Young, *Large deviations in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **28** (2008), no. 2, 587–612.

[Ryc83] Marek Rychlik, *Bounded variation and invariant measures*, Studia Math. **76** (1983), no. 1, 69–80.

[Sch15] Tanja Schindler, *Generalized strong laws of large numbers for intermediately trimmed sums for non-negative stationary processes*, Ph.D. thesis, Mathematik and Informatik, Universität Bremen, 2015, `http://elib.suub.uni-bremen.de/edocs/00104900-1.pdf`.

[Tak84] Y. Takahashi, *Entropy functional (free energy) for dynamical systems and their random perturbations*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 437–467.

[Tak87] ______, *Asymptotic behaviours of measures of small tubes: entropy, Liapunov's exponent and large deviation*, Dynamical systems and applications (Kyoto, 1987), World Sci. Adv. Ser. Dynam. Systems, vol. 5, World Sci. Publishing, Singapore, 1987, pp. 1–21.

[Var84] S. R. S. Varadhan, *Large deviations and applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 46, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984.

[Var08] ______, *Large deviations*, Ann. Probab. **36** (2008), no. 2, 397–419, `https://arxiv.org/abs/0804.2330v1`.

[Var12] Paulo Varandas, *Non-uniform specification and large deviations for weak Gibbs measures*, J. Stat. Phys. **146** (2012), no. 2, 330–358.

[Wad96] Simon Waddington, *Large deviation asymptotics for Anosov flows*, Ann. Inst. H. Poincaré Anal. Non Linéaire **13** (1996), no. 4, 445–484.

[You90] Lai-Sang Young, *Large deviations in dynamical systems*, Trans. Amer. Math. Soc. **318** (1990), no. 2, 525–543.

[You98] ______, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650.

[You99] ______, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188.

[Zwe98] Roland Zweimüller, *Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points*, Nonlinearity **11** (1998), no. 5, 1263–1276.