Erdős-Rényi laws for exponentially and polynomially mixing dynamical systems.

Nicolai Haydn and Matthew Nicol *

April 13, 2023

Abstract

Erdős-Rényi limit laws give the length scale of a time-window over which time-averages in Birkhoff sums have a non-trivial almost-sure limit. We establish Erdős-Rényi type limit laws for Hölder observables on dynamical systems modeled by Young Towers with exponential and polynomial tails. This extends earlier results on Erdős-Rényi limit laws to a broad class of dynamical systems with some degree of hyperbolicity.

1 Introduction

The Erdős-Rényi fluctuation law gives the length scale of a time-window over which time-averages in Birkhoff sums have a non-trivial almost-sure limit. It was first proved in the independent and identically distributed (i.i.d.) case [11] in the following form:

Proposition 1.1. *Let $(X_n)_{n\geq 1}$ be an i.i.d. sequence of non-degenerate random variables, $\mathbb{E}[X_1] = 0$, and let $S_n = X_1 + \cdots + X_n$. Assume that the moment generating function $\phi(t) = \mathbb{E}(e^{tX_1})$ exists in some open interval $U \subset \mathbb{R}$ containing $t = 0$. For each $\alpha > 0$, define $\psi_\alpha(t) = \phi(t)e^{-\alpha t}$. For those α for which ψ_α attains its minimum at a point $t_\alpha \in U$, let $c_\alpha = \alpha t_\alpha - \ln \phi(t_\alpha)$. Then*

$$\lim_{n\to\infty} \max\{(S_{j+[\ln n/c_\alpha]} - S_j)/[\ln n/c_\alpha] : 1 \leq j \leq n - [\ln n/c_\alpha]\} = \alpha$$

In the theorem above the Gauss bracket $[\cdot]$ denotes the integer part of a number. The existence of $\psi_\alpha(t)$ for all $t \in U$ implies exponential large deviations with a rate function (in fact $c_\alpha = I(\alpha)$ where I is the rate function, defined later) and this implies that sampling over a window length $k(n)$ of larger than logarithmic length scale (in the sense that $k(n)/\ln n \to \infty$), allows the ergodic theorem to kick in and

$$\lim_{n\to\infty} \max\{(S_{j+k(n)} - S_j)/k(n) : 1 \leq j \leq n - k(n)\} = 0$$

while sampling over too small a window, for example $k(n) = 1$, gives similarly a trivial limit

$$\lim_{n\to\infty} \max\{(S_{j+k(n)} - S_j)/k(n) : 1 \leq j \leq n - k(n)\} = \|X_1\|_\infty$$

Define the function

$$\theta(n, k(n)) := \max_{0\leq j\leq n-k(n)} S_{j+k(n)} - S_j,$$

*Department of Mathematics, University of Southern California; Department of Mathematics, University of Houston. E-mail: <nhaydn@usc.edu>,<nicol@math.uh.edu>. MN would like to thank the NSF for support on NSF-DMS Grant 2009923.

which may be interpreted as the maximal average gain over a time window of length $k(n)$ up to time n. In the setting of coin tosses the Erdős-Rényi law gives precise information on the maximal average gain of a player in a fair game in the case where the length of the time window ensures $\lim_{n\to\infty} \frac{\theta(n,k(n))}{k(n)}$ has a non-degenerate almost sure limit.

In 1986 Deheuvels, Devroye and Lynch [7] in the i.i.d. setting of Proposition 1.1 gave a precise rate of convergence and showed that if $k(n) = [\ln n/c_\alpha]$ then P a.s:

$$\limsup \frac{[\theta(n,k(n)) - \alpha k(n)]}{\ln k(n)} = \frac{1}{2t_\alpha}$$

and

$$\liminf \frac{[\theta(n,k(n)) - \alpha k(n)]}{\ln k(n)} = -\frac{1}{2t_\alpha}$$

In this paper we establish Erdős-Rényi limit laws for Hölder observables on dynamical systems modeled by Young Towers [23, 24] with exponential and polynomial tails (see section 2). Tails refer to the measure $\mu(R > n)$ of the return time R function to the base of the tower. Our exposition is based upon [15, Section 2.3] and [17] who present a framework more general than that of the original Tower construction of Young [23] in that uniform contraction of local stable manifolds is not assumed for polynomially mixing systems in dimensions greater than 1. We will give more details on Young Towers below but here note that Hölder observables on Young Towers with exponential (polynomial) tails have exponential (polynomial) decay of correlations, the precise rate is encoded in the return time function.

Our results extend the work of [18] from the class of non-uniformly expanding maps with exponential decay of correlations to all systems modeled by a Young Tower, including Sinai dispersing billiard maps; diffeomorphisms of Henón type; polynomially mixing billards as in [4] (as long as the correlation decay rate is greater than $n^{-\beta}$, $\beta > 1$); smooth unimodal and multimodal maps satisfying the Collet-Eckmann conditions [15, Example 4.10]; certain Viana maps [15, Example 4.11]; and Lorenz-like maps. Other examples to which our results apply are listed in [17].

In the setting of hyperbolic dynamical systems there are many earlier results. Grigull [12] established the Erdős-Rényi law for hyperbolic rational maps, Chazottes and Collet [5] proved Erdős-Rényi theorems with rates for uniformly expanding maps of the interval, while Denker and Kabluchko [8] proved Erdős-Rényi results for Gibbs-Markov dynamics. In [9] Erdős-Rényi limit laws for Lipschitz observations on a class of non-uniformly expanding dynamical systems, including logistic-like maps, were given as well as related results on maximal averages of a time series arising from Hölder observations on intermittent-type maps over a time window of polynomial length. Kifer [13, 14] has established Erdős-Rényi laws for non-conventional ergodic sums and in the setting of averaging or homogenization of chaotic dynamical systems. We mention also recent related work of [2, 3] on applications of Erdő-Rényi limit laws to multifractal analysis.

The main novelty of our technique is the use of the symbolic metric on the axiomatic Young Tower construction of [17, 15] to control the norm of the indicator function of sets of the form $(S_n > n\alpha)$ on the quotiented tower. This eliminates many difficulties involved with considering the Lipschitz norm of such sets with respect to the Riemannian metric on the phase space of the system. The structure allows us to consider, with small error, averaged Birkhoff sums as being constant on stable manifolds, and thence use the decay of correlations for observables on the quotiented tower in terms of their Lipschitz and L^∞ norms.

Our results in the case of Young Towers with exponential decay of correlations, Theorem 5.1, are optimal and replicate the i.i.d case, while in the case of Young Towers with polynomial tails we investigate windows of polynomial length and notice that there is still a gap between upper and lower bounds, Theorem 7.1 and Theorem 7.2 which however can be quite small as Example 7.3 shows.

2 Young Towers.

We now describe more precisely what we mean by a non-uniformly hyperbolic dynamical system modeled by a Young Tower. Our exposition is based upon [15, Section 2.3] and [17] who present a framework more general than that of the original Young Tower [23] in that uniform contraction of local stable manifolds is not assumed for polynomially mixing systems in dimensions greater than 1. This set-up is very useful for the study of almost sure fluctuations of Birkhoff sums of bounded variables.

We suppose T is a diffeomorphism of a Riemannian manifold (M, d), possibly with singularities. Fix a subset $\Lambda \subset M$ with a 'product structure'. Product structure means there exists a family of disjoint stable disks (local stable manifolds) $\{W^s\}$ that cover Λ as well as a family of disjoint unstable disks (local unstable manifolds) $\{W^u\}$ that cover Λ. The stable and unstable disks containing $x \in \Lambda$ are denoted $W^s(x)$ and $W^u(x)$. Each stable disk intersects each unstable disk in precisely one point.

Suppose there is a partition $\{\Lambda_j\}$ of Λ such that each stable disk $W^s(x)$ lies in Λ_j if $x \in \Lambda_j$. Suppose there exists a 'return time' integer-valued function $R : \Lambda \to \mathbb{N}$, constant with value $R(j)$ on each partition element Λ_j, such that $T^{R(j)}(W^s(x)) \subset W^s(T^{R(j)}x)$ for all $x \in \Lambda_j$. We assume that the greatest common denominator of the integers $\{R(j)\}$ is 1, which ensures that the Tower is mixing. We define the induced return map $f : \Lambda \to \Lambda$ by $f(x) = T^{R(x)}(x)$.

For $x, y \in \Lambda$ let $s(x, y)$ be the least integer $n \geq 0$ such that $f^n(x)$ and $f^n(y)$ lie in different partition elements of Λ. We call s the separation time with respect to the map $f : \Lambda \to \Lambda$.
Assumptions: there exist constants $K \geq 1$ and $0 < \beta_1 < 1$ such that

(a) if $z \in W^s(x)$ then $d(f^n z, f^n x) \leq K\beta_1^n$;

(b) if $z \in W^u(x)$ then $d(f^n z, f^n x) \leq K\beta_1^{s(x,z)-n}$;

(c) if $z, x \in \Lambda$ then $d(T^j z, T^j x) \leq K(d(z, x) + d(fz, fx))$ for all $0 \leq j \leq \min\{R(z), R(x)\}$.

Define an equivalence relation on Λ by $z \sim x$ if $z \in W^s(x)$ and form the quotient space $\overline{\Lambda} = \Lambda/\sim$ with corresponding partition $\{\overline{\Lambda_j}\}$. The return time function $R : \overline{\Lambda} \to \mathbb{N}$ is well-defined as each stable disk $W^s(x)$ lies in Λ_j if $x \in \Lambda_j$ and $T^{R(j)}(W^s(x)) \subset W^s(T^{R(j)}x)$ for all $x \in \Lambda_j$. This defines the induced map $\bar{f} : \overline{\Lambda} \to \overline{\Lambda}$. Suppose that $\bar{f}$ and the partition $\{\overline{\Lambda_j}\}$ separates points in $\overline{\Lambda}$. Define $d_{\beta_1}(z, x) = \beta_1^{s(z,x)}$, then d_{β_1} is a metric on $\overline{\Lambda}$.

Let m be a reference probability measure on $\overline{\Lambda}$ (in most applications this will be normalized Lebesgue measure). Assume $\int R\, dm < \infty$ and that $\bar{f} : \overline{\Lambda} \to \overline{\Lambda}$ is Gibbs-Markov uniformly expanding on $(\overline{\Lambda}, d_{\beta_1})$. By this we mean that $\bar{f}$ is a measure-theoretic bijection from each $\overline{\Lambda_j}$ onto $\overline{\Lambda}$.

We assume that $\bar{f} : \overline{\Lambda} \to \overline{\Lambda}$ has an invariant probability measure $\bar{\nu}$ and $0 < a < \frac{d\bar{\nu}}{dm} < b$ for some constants a, b. Then R is $\bar{\nu}$-integrable and there is an f invariant probability ν measure on Λ such that $\bar{\pi}^*\nu = \bar{\nu}$ where $\bar{\pi}$ is the quotient map taking Λ onto $\Lambda/\sim$.

Now we define the Young Tower

$$\Delta = \{(x, j) \in \Lambda \times \mathbb{N} : 0 \leq j \leq R(x) - 1\}$$

and the tower map F by

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

We extend the definition of the return time function R to Δ by defining $R(x,j) = R(x) - j$. We lift ν in a standard way to an invariant probability measure ν_Δ for $F:\Delta\to\Delta$. In fact $\nu_\Delta = \nu\times$ counting measure. The separation function s can be extended to the full Young Tower Δ by

$$s((x,l),(y,l^{'})) = \begin{cases} s(x,y) & \text{if } l = l^{'}; \\ 1 & \text{if } l \neq l^{'}. \end{cases}$$

Define the semi-conjugacy $\pi:\Delta\to M$, $\pi(x,j) = T^j(x)$ for $0\le j < R(x)$. The measure $\mu = \pi^*\nu_\Delta$ is a T-invariant mixing probability measure on M. Given an observable $\varphi: M\to\mathbb{R}$, we may lift to an observable $\varphi:\Delta\to\mathbb{R}$ by defining $\varphi(x,j) = \varphi(T^jx)$ (we keep the same notation for the observable). The semi-conjugacy π^* allows us to transfer statistical properties from lifted observables φ on (Δ, F, ν_Δ) to the original observables φ on (T, M, μ) [23].

We now define the metric d_{β_1} on $\overline{\Delta}$ by $d_{\beta_1}(p,q) = \beta_1^{s(p,q)}$. Here we write, for convenience, $p = (x,l)\in\overline{\Delta}$, $q = (y,l')\in\overline{\Delta}$. We define the $\|\cdot\|_{\beta_1}$-norm by $\|\phi\|_{\beta_1} := \|\phi\|_\infty + \sup_{p,q\in\Delta}\frac{|\phi(p)-\phi(q)|}{d_{\beta_1}(p,q)}$. Functions ϕ and ψ which are constant on stable manifolds in Δ naturally project to functions ϕ and ψ (we use the same notation) on $\overline{\Delta}$ with the same d_{β_1} Lipschitz constant and L^∞ norm. If $\phi:\Delta\to\mathbb{R}$ is constant on stable manifolds we define the $\|\cdot\|_{\beta_1}$-norm by $\|\phi\|_{\beta_1} := \|\phi\|_\infty + \sup_{p,q\in\Delta}\frac{|\phi(p)-\phi(q)|}{d_{\beta_1}(p,q)}$.

In the proof of Theorem 5.1 the following result will be useful.

Proposition 2.1. *[15, Corollary 2.9] Suppose that ϕ, $\psi:\Delta\to\mathbb{R}$ are constant on stable manifolds and $\|\phi\|_{\beta_1}<\infty$ and $\|\psi\|_\infty<\infty$.*

Then there exist constants C, $\beta_3\in(0,1)$ so that

$$\left|\int_\Delta \phi(\psi\circ F^j)\,d\nu_\Delta - \int_\Delta \phi\,d\nu_\Delta \int_\Delta \psi\,d\nu_\Delta\right| \le C\|\phi\|_{\beta_1}\|\psi\|_\infty\beta_3^j$$

for all $j\ge 0$.

This result is also implied by [23].Since we assume that the function ψ is constant on loccal stable leaves, we only need it to be bounded and not Lipschitz continuous since this does not require the approximation argument of section 4.

3 Large deviations and rate functions.

Before stating precisely our main result we recall the definition of rate function and some other notions of large deviations theory. Suppose (T, M, μ) is a probability preserving transformation and $\varphi: M\to\mathbb{R}$ is a mean-zero integrable function i.e. $\int_M\varphi\,d\mu = 0$. 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. The subject of large deviations is concerned with the probability of deviation of $\frac{1}{n}S_n(\varphi)$ from its asymptotic limit of zero, in particular for $\alpha > 0$

$$\mu\{x\in M : \frac{1}{n}S_n(\varphi)(x)\ge\alpha\}$$

and if $\alpha < 0$

$$\mu\{x\in M : \frac{1}{n}S_n(\varphi)(x)\le\alpha\}$$

If these quantities tend to zero at an exponential rate then sometimes the rate is determined by a *rate function* $I(\alpha)$ (see [10, Chapter 2 Section 2.6] for a discussion of rate functions in the iid case and [18, 21] in the deterministic case).

Definition 3.1. *A mean-zero integrable function $\varphi : M \to \mathbb{R}$ 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*

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

for all $\alpha > 0$ in U and

$$\lim_{n\to\infty} \frac{1}{n} \log \mu(S_n(\varphi) \leq n\alpha) \quad = \quad -I(\alpha) \tag{2}$$

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.

For Hölder observables on Young Towers with exponential tails (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 $I_\varphi(\alpha)$ [18, 21, 16, 20]. A formula for the width of U is given in [21] following a standard approach but it is not useful in concrete estimates.

4 Erdős-Rényi laws: background.

Proposition 4.1 given below is found in the form we use in [9], where a proof is also given. It is due to Erdős and Rényi [11] (for variants see also [6, Theorem 2], Grigull [12], or Denker and Kabluchko [8]). Recall that the Gauss bracket $[\cdot]$ 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 4.1. *[9] Let (T, M, μ) be an ergodic dynamical system and $\varphi : M \to \mathbb{R}$ an observable.*

(a) Suppose that φ satisfies a large deviation principle with rate function I defined on the open set U and assume $\mu(\varphi) = 0$. Let $\alpha > 0$, $\alpha \in U$ and set

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

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

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

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

$$\mu\left(\bigcap_{m=0}^{n-L_n} \{S_{L_n}(\varphi) \circ T^m \in A\}\right) \quad \leq \quad C[\mu(S_{L_n} \in A)]^{n/(L_n)^\kappa} \tag{3}$$

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

$$\liminf_{n\to\infty} \max_{0\leq j\leq n-L_n} \frac{1}{L_n} S_{L_n}(\varphi) \circ T^j \geq \alpha.$$

Remark 4.2. *If both Assumptions (a) and (b) of Proposition 4.1 hold then*

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

Remark 4.3. *The proof of Proposition 4.1 shows that the upper Erdős-Rényi law follows from the existence of exponential large deviations given by a rate function, while for the lower Erdős-Rényi law it suffices to show that for every* $\epsilon > 0$ *the series* $\sum_{n>0} \mu(B_n(\epsilon))$, *where* $B_n(\epsilon) = \{\max_{0\le m\le n-L_n} S_{L_n} \circ T^m \le L_n(\alpha-\epsilon)\}$ *is summable. This is usually the harder part to prove in the deterministic case.*

5 Erdős-Rényi limit laws for Young Towers with exponential tails.

We now state our main theorem in the case of exponential tails.

Theorem 5.1. *Suppose* (T, M, μ) *is a dynamical system modeled by a Young Tower with* $\nu_\Delta(R > j) \le C_2\beta_2^j$ *for some* $\beta_2 \in (0,1)$ *and some constant* C_2. *Let* $\varphi : M \to \mathbb{R}$ *be Hölder with* $\int \varphi \, d\mu = 0$. *Assume* $\varphi \neq \psi\circ T - \psi$ *for any* $\psi \in L^1(\mu)$. *Let* $I(\alpha)$ *denote the rate function defined on an open set* $U \subset \mathbb{R}$ *containing* 0.

Let $\alpha > 0$, $\alpha \in U$ *and define*

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

Then

$$\lim_{n\to\infty} \max_{0\le j\le n-L_n} \frac{S_{L_n} \circ T^j(x)}{L_n} = \alpha,$$

for μ *a.e.* $x \in \Omega$, *where as before* $S_m(x) = \sum_{j=0}^{m-1} \varphi(T^j x)$.

6 Proof of Theorem 5.1.

We will prove that we have an upper bound as in Assumption (a) of Proposition 4.1 and a lower bound as in Assumption (b) of Proposition 4.1. If we obtain the upper and lower bound then by Remark 4.2 we obtain the limit we wish.

6.1 Upper bound.

The upper bound is straightforward. In the case that φ is not an L^1 coboundary i.e. there exists no ψ such that $\varphi = \psi \circ T - \psi$, $\psi \in L^1(m)$ it has been shown [18, 21] under the assumptions of Theorem 5.1 that φ has exponential large deviations with a rate function $I(\alpha)$. Thus assumption (a) of Proposition 4.1 holds and we automatically have an upper bound.

6.2 Lower bound.

In order to obtain the lower bound, by Remark 4.3, we only need to prove $\mu(\{\max_{0\le m\le n-L_n} S_{L_n} \circ T^m \le L_n(\alpha-\epsilon)\})$ is summable. This direction is the more difficult part of the proof and uses differential and dynamical information on the system.

Throughout this proof we will assume that φ is Lipschitz, as the modification for Hölder φ is straightforward.

The next lemma is not optimal but is useful in allowing us to go from uniform contraction along stable manifolds upon returns to the base of the Young Tower (Property (P3) of [23]) to estimates of the contraction along stable leaves in the whole manifold.

Lemma 6.1. *Let β_1 be defined as in Section (2.1) Assumption (a) and β_2 as in Theorem 5.1. Let $D(m) = \{(x,j) \in \Delta : |T^k W^s(x,j)| < 2K\beta_1^{\sqrt{k}}$ for all $k \geq m\}$. Then for any $\delta > 0$ there exists $N(\delta) > 0$ such that for all $m \geq N(\delta)$, $\nu_\Delta(D(m)^c) \leq (\beta_2 + \delta)^{\sqrt{m}}$.*

Proof. Let $\tau_r(x,j) := \#\{k : 1 < k \leq r : F^k(x,j) \in \Lambda\}$ denote the number of times $k \in [1,r]$ that $F^k(x,j)$ lies in the base Λ of the Young Tower. Let $B_r = \{(x,j) \in \Delta : \tau_r(x,j) \leq \sqrt{r}\}$. If $(x,j) \in B_r$ then $\tau_r(x,j) \leq \sqrt{r}$ and there is at least one $k \in [0,r]$, such that $R(F^k(x,j)) > \sqrt{r}$. Hence $B_r \subset \bigcup_{k=1}^{r} F^{-k}(R > \sqrt{r})$ and thus by assumption of exponential tails: $\nu_\Delta(B_r) \leq r\nu(R > \sqrt{r}) < Cr{\beta_2}^{\sqrt{r}}$.

Suppose now that $(x,j) \in B_r^c$. Then $|T^r W^s((x,j))| \leq 2K\beta_1^{\sqrt{r}}$ by assumptions (a) and (c). Now $\nu_\Delta(\bigcup_{r\geq m} B_r) \leq \sum_{r\geq m} Cr{\beta_2}^{\sqrt{r}} \leq (\beta_2+\delta)^{\sqrt{m}}$ for all m large enough. The lemma now follows as $D(m)^c \subset \bigcup_{r=m}^{\infty} B_r$ and so $\nu_\Delta(D(m)^c) \leq (\beta_2+\delta)^{\sqrt{m}}$ □

Corollary 6.2. *Lift $\varphi : M \to \mathbb{R}$ to $\varphi : \Delta \to \mathbb{R}$ by defining $\varphi(x,j) = \varphi(T^j x)$. Let β_1 be defined as in Section (2.1) Assumption (a). Suppose $p \in D(m) = \{(x,j) \in \Delta : |T^k W^s(x,j)| < 2K\beta_1^{\sqrt{k}}$ for all $k \geq m\}$ and let $L_n = [\frac{\ln n}{I(\alpha)}]$. Then if $q \in W^s(p)$, $|S_{L_n}\varphi \circ F^m(p) - S_{L_n}\varphi \circ F^m(q)| \leq C\|\varphi\|_{Lip} L_n {\beta_1}^{\sqrt{m}}$.*

Proof. By Lemma 6.1

$$
\begin{aligned}
|S_{L_n}\varphi \circ F^m(p) - S_{L_n}\varphi \circ F^m(q)| &\leq \sum_{j=0}^{L_n-1} \|\varphi\|_{Lip} |T^{j+m}(p) - T^{j+m}(q)| \\
&\leq \|\varphi\|_{Lip} \sum_{j=0}^{L_n-1} 2K\beta_1^{\sqrt{j+m}} \\
&\leq 2K\|\varphi\|_{Lip} L_n \beta_1^{\sqrt{m}}
\end{aligned}
$$

which proves the corollary with $C = 2K$. □

Proof of Theorem 5.1. The main idea of the proof of Theorem 5.1 is to approximate functions on Δ by functions constant on stable manifolds, so that correlation decay estimates on the quotiented tower from Proposition 2.1 can be used.

We lift φ from M to Δ by defining $\varphi(x,j) = \varphi(T^j x)$. We will use the same notation for φ on Δ as we use for φ on M.

To simplify notation we will sometimes write $p = (x,j) \in \Delta$ for points in Δ. For $0 < \epsilon \ll \alpha$ put

$$A_n(\epsilon) := \{(x,j) \in \Delta : S_{L_n} \leq L_n(\alpha - \epsilon)\},$$

where

$$S_n(x,j) = \sum_{k=0}^{n-1} \varphi \circ F^k(x,j)$$

is the nth ergodic sum of φ. Define

$$B_n(\epsilon) = \bigcap_{m=0}^{n-L_n} F^{-m} A_n(\epsilon) = \left\{(x,j) \in \Delta : \max_{0\leq m\leq n-L_n} S_{l_n} \circ F^m \leq L_n(\alpha - \epsilon)\right\}.$$

The theorem follows by the Borel-Cantelli lemma once we show that $\sum_{n=1}^{\infty} \nu_\Delta(B_n(\epsilon)) < \infty$. To do this we will use a blocking argument to take advantage of decay of correlations and intercalate by blocks of length $\kappa_n := \ln^\kappa n$ for some $\kappa > 2$ (which turns out to be sufficient).

For $1 \leq j < r_n := [\frac{n}{\kappa_n}]$ put

$$E_n^j(\epsilon) := \bigcap_{m=0}^{j} F^{-m[\kappa_n]} A_n(\epsilon)$$

which for every n is a nested sequence of sets, that is $E_n^{j+1} \subset E_n^j$. Then $B_n(\epsilon) \subset E_n^{r_n}(\epsilon)$ and $\nu_\Delta(B_n(\epsilon)) \leq \nu_\Delta(E_n^{r_n}(\epsilon))$. We also have the recursion

$$E_n^j(\epsilon) = A_n(\epsilon) \cap F^{-\kappa_n} E_n^{j-1}(\epsilon)$$

$j = 1, \ldots, r_n$, which implies

$$\nu_\Delta(E_n^j(\epsilon)) = \nu_\Delta(A_n(\epsilon) \cap F^{-\kappa_n} E_n^{j-1}(\epsilon))$$

Let $D(m) = \{(x, j) \in \Delta : |T^k W^s(x)| < 2K\beta_1^{\sqrt{k}}$ for all $k \geq m\}$ as in Lemma 6.1. Hence, given $\delta > 0$ such that $\beta_2' := \beta_2 + \delta < 1$ then by Lemma 6.1 $\nu_\Delta(D(\kappa_n)^c) \leq \beta_2'^{\sqrt{\kappa_n}}$ for sufficiently large n.

Furthermore, if $m \geq \kappa_n$ then $D(\kappa_n) \subset D(m)$ and if $p \in D(m)$ and $q \in W^s(p)$ then by Corollary 6.2

$$|S_{L_n} \circ F^m(p) - S_{L_n} \circ F^m(q)| \leq C\|\varphi\|_{Lip} L_n \beta_1^{\sqrt{m}}$$

Since $m > \ln^2(n)$, $C\|\varphi\|_{Lip} L_n \beta_1^{\sqrt{m}} \to 0$ as $n \to \infty$. Thus $|S_{L_n} \circ F^m(p) - S_{L_n} \circ F^m(q)| < \epsilon/2$ for all n large enough.

Accordingly for large n if $m \geq \kappa_n$, $p = (x, j) \in D(m) \cap F^{-m} A_n(\epsilon)$ and $q \in W^s(p)$ then $F^m q \in A_n(\frac{\epsilon}{2})$.

Approximation by functions constant on local stable leaves.

We now approximate $1_{F^{-\kappa_n} A_n(\epsilon) \cap D(\kappa_n)}$ by a function g_n^ϵ which is constant on stable manifolds by requiring that if $p \in F^{-\kappa_n} A_n(\epsilon) \cap D(\kappa_n)$ then $g_n^\epsilon(p) = 1$ on $W^s(p)$ and $g_n^\epsilon = 0$ otherwise. Thus $\{g_n^\epsilon = 1\} \subset A_n(\frac{\epsilon}{2})$ and

$$\nu_\Delta(g_n^\epsilon = 1) \leq \nu_\Delta(A_n(\frac{\epsilon}{2}))$$

Furthermore

$$F^{-\kappa_n} A_n(\epsilon) \subset \{g_n^\epsilon = 1\} \cup D(\kappa_n)^c$$

and hence

$$\nu_\Delta(A_n(\epsilon)) \leq \nu_\Delta(g_n^\epsilon = 1) + \nu_\Delta(D(\kappa_n)^c).$$

For $j = 1, \ldots, r_n$ let

$$G_n^j(\epsilon) =: \prod_{i=0}^{j} g_n^\epsilon \circ F^{i[\kappa_n]}$$

and note $\nu_\Delta(E_n^j(\epsilon)) \leq \nu_\Delta(G_n^j(\epsilon)) + j\nu_\Delta(D(\kappa_n)^c)$.

Smoothing the approximation functions constant on local stable leaves.

We will approximate g_n^ϵ (considered as a function on $\overline{\Delta}$) by a d_{β_1} Lipschitz function h_n^ϵ which extends to a function on Δ and is also constant on stable leaves.

First define

$$h_n^\epsilon(\bar{p}) := \max\{0, 1 - d_{\beta_1}(\bar{p}, \mathrm{supp}(g_n^\epsilon))\beta_1^{-\sqrt{\kappa_n}}\}$$

on $\overline{\Delta}$ and then extend so that it is constant on local stable manifolds and hence is a function on Δ. In particular h_n^ϵ has support in points such that $d_{\beta_1}(p, \mathrm{supp}(g_n^\epsilon)) \le \beta_1^{\sqrt{\kappa_n}}$ and $\|h_n^\epsilon\|_{\beta_1} \le \beta_1^{\sqrt{\kappa_n}}$ by [22, Section 2.1].

By Section 2 Assumptions (b) and (c) if $z \in W^u(p)$ and $d_{\beta_1}(p,z) < \beta_1^{\sqrt{\kappa_n}}$ then $d(F^j p, F^j z) \le 2K\beta_1^{\sqrt{\kappa_n}-L_n}$ for all $j \le L_n$ provided $L_n \le \sqrt{\kappa_n}$.

Hence if $d_{\beta_1}(z, \mathrm{supp}(g_n^\epsilon)) \le \beta_1^{\sqrt{\kappa_n}}$ then there exists $p \in \mathrm{supp}(g_n^\epsilon)$ such that $d(F^j p, F^j z) \le 2K\beta_1^{\sqrt{\kappa_n}-L_n}$ for all $j \le L_n$ and hence

$$|\sum_{j=0}^{L_n}[\varphi \circ F^j(z) - \varphi \circ F^j(p)]| \le \|\varphi\|_{Lip} \sum_{j=0}^{L_n} d(F^j p, F^j q) \le C L_n \beta_1^{\sqrt{\kappa_n}-L_n} \le \frac{\epsilon}{2}$$

for all sufficiently large n. This implies that $\nu_\Delta(g_n^\epsilon) \le \nu_\Delta(h_n^\epsilon) \le \nu_\Delta(A_n(\frac{\epsilon}{2}))$.

As h_n^ϵ Lipschitz in the d_{β_1} metric we obtain by Proposition 6.3 with $\beta_1' \in (\beta_2, 1)$

$$\begin{aligned}
\nu_\Delta(E_n^j(\epsilon)) &\le \int_\Delta (G_n^j(\epsilon))\, d\nu_\Delta + j\nu_\Delta(D(\kappa_n)^c) \\
&\le \int_\Delta (g_n^\epsilon \cdot G_n^{j-1} \circ F^{\kappa_n})\, d\nu_\Delta + Cj{\beta_2'}^{\sqrt{\kappa_n}} \\
&\le \int h_n^\epsilon\, d\nu_\Delta \int G_n^{j-1}(\epsilon)\, d\nu_\Delta + c_3\beta_3^{\kappa_n}\|h_n^\epsilon\|_{\beta_1}\|G_n^{j-1}(\epsilon)\|_\infty + Cj{\beta_2'}^{\sqrt{\kappa_n}} \\
&\le \nu_\Delta(A_n(\frac{\epsilon}{2}))\nu_\Delta(G_n^{j-1}(\epsilon)) + c_3\beta_3^{\kappa_n}\beta_1^{-\sqrt{\kappa_n}} + Cj{\beta_2'}^{\sqrt{\kappa_n}}.
\end{aligned}$$

Iterating the estimate for $\int_\Delta (G_n^j(\epsilon))\, d\nu_\Delta$ yields

$$\nu_\Delta(E_n^{r_n}(\epsilon)) \le \nu_\Delta(A_n(\frac{\epsilon}{2}))^{[n/\kappa_n]} + nc_3\beta_3^{\kappa_n}\beta_1^{-\sqrt{\kappa_n}} + n^2 C{\beta'}_2^{\sqrt{\kappa_n}}.$$

The terms $nc_3\beta_3^{\kappa_n}\beta_1^{-\sqrt{\kappa_n}}$ and $n^2 C{\beta_2'}^{\sqrt{\kappa_n}}$ are summable since $\kappa > 2$.

Using the properties of the rate function.

In order to verify summability of the principal terms $\nu_\Delta(A_n(\frac{\epsilon}{2}))^{[n/\kappa_n]}$ term we proceed as in the proof of Proposition 4.1 using the existence of a large deviations rate function. By the strict convexity of the rate function on a neighborhood U of zero and the fact that $I(0) = 0$ we obtain $\nu_\Delta((A_n(\frac{\epsilon}{2}))^c) \ge e^{-L_n(I(\alpha - \frac{\epsilon}{2})+\delta_1)}$ for some $0 < \delta_1$ and hence $1 - \nu_\Delta(A_n(\frac{\epsilon}{2})) \ge e^{-L_n(I(\alpha-\frac{\epsilon}{2})+\delta_1)}$ for some $0 < \delta_1$. Hence $\nu_\Delta(A_n(\frac{\epsilon}{2})) \le 1 - n^{-\rho}$ where $\rho = \frac{I(\alpha-\frac{\epsilon}{2})}{I(\alpha)} + \delta_1$ is less than 1 for $\delta_1 > 0$ small enough. The principal term can be bounded by

$$\nu_\Delta(A_n(\frac{\epsilon}{2}))^{[n/\kappa_n]} \le (1 - n^{-\rho})^{[n/\kappa_n]}$$

which is also summable over n. Hence by the Borel-Cantelli lemma we conclude that the set $\{B_n(\epsilon) \text{ i.o.}\}$ has measure zero. This concludes the proof. □

7 Erdös-Rényi laws for Young Towers with polynomial tails.

We now consider Young Towers with polynomial tails in the sense that $\nu_\Delta(R > n) \le Cn^{-\beta}$, $\beta > 1$. It is shown in [18, Theorem 3.1, Theorem 4.2] that this implies that if φ is a Hölder observable on (T, M, μ) modeled by such a tower then for all $\delta > 0$

$$\mu\left(\left|\frac{1}{n}S_n(\varphi) - \bar{\varphi}\right| > \epsilon\right) \le C(\epsilon)n^{-\beta-\delta}.$$

These results are extended in [16] to the setting $\beta > 0$. Lower bounds are also given in [18, Proposition 3.3, Theorem 3.5] and [16, Proposition A.1, Corollary A.2] which show that this large deviation rate is close to optimal.

7.1 Upper bounds.

We first prove a general result. We suppose that (T, M, μ) is an ergodic dynamical system and $\varphi : M \to \mathbb{R}$ is a bounded observable. We assume also that there exists $\beta > 1$ such that for all $\epsilon > 0$ and for all $n \geq 0$ there exists a constant $C(\epsilon)$ such that

$$\mu\left(\left|\frac{1}{n}S_n(\varphi) - \bar{\varphi}\right| > \epsilon\right) \leq C(\epsilon)n^{-\beta}.$$

Theorem 7.1. *Assume that $\bar{\varphi} = \mu(\varphi) = 0$, φ is bounded, and that there exists $\beta > 1$ such that for every $\epsilon > 0$ there exists a constant $C(\epsilon) > 0$ so that*

$$\mu\left(\left|\frac{1}{n}S_n(\varphi)\right| > \epsilon\right) \leq C(\epsilon)n^{-\beta}.$$

Then if $\tau \in (\frac{1}{\beta}, 1)$ for μ a.e. $x \in M$,

$$\lim_{n\to\infty} \max_{0\leq m\leq n-n^\tau} n^{-\tau} S_{n^\tau} \circ T^m(x) = 0.$$

Proof. Choose $\tau \in (\frac{1}{\beta}, 1)$ and put $L_n = n^\tau$. Let $\epsilon > 0$ and define

$$A_n := \left\{x \in M : \max_{0\leq m\leq n-L_n} |S_{L_n} \circ T^m| \geq L_n\epsilon\right\}.$$

Then $\mu(A_n) \leq n\mu(S_{L_n} \geq \epsilon L_n) \leq C(\epsilon)n^{1-\tau\beta} = C(\epsilon)n^{-\delta}$, where $\delta = \tau\beta - 1$.

Let $p > \frac{1}{\delta}$ (i.e. $\delta p > 1$) and consider the subsequence $n = k^p$. Since $\sum_k \mu(A_{k^p}) \leq C(\epsilon)\sum_k k^{-p\delta} < \infty$, we obtain via the Borel-Cantelli lemma that for μ a.e. $x \in M$

$$\limsup_{k\to\infty} \max_{0\leq m\leq k^p-L_{k^p}} L_{k^p}^{-1}|S_{L_{k^p}} \circ T^m| \leq \epsilon.$$

To fill the gaps use that $k^p - (k-1)^p = O(k^{p-1})$ and we obtain (as φ is bounded) that

$$\frac{S_{L_{k^p}} \circ T^m}{L_{k^p}} = \frac{S_{L_{(k-1)^p}} \circ T^m}{L_{k^p}} + \mathcal{O}\left(\frac{1}{k}\right)$$

where the implied constant is uniform in $x \in M$ as φ is bounded. As $\lim_{k\to\infty} \frac{k^p}{(k-1)^p} = 1$ we conclude

$$\limsup_{k\to\infty} \max_{0\leq m\leq k^p-L_{k^p}} \frac{|S_{L_{k^p}} \circ T^m|}{L_{k^p}} = \limsup_{k\to\infty} \max_{0\leq m\leq (k-1)^p-L_{(k-1)^p}} \frac{|S_{L_{(k-1)^p}} \circ T^m|}{L_{k^p}}.$$

Since any $n \in \mathbb{N}$ satisfies $(k-1)^p \leq n \leq k^p$ for some k and φ is bounded, it follows that

$$\limsup_{n\to\infty} \max_{0\leq m\leq n-L_n} |S_{L_n} \circ T^m|/L_n \leq \epsilon.$$

As ϵ was arbitrary this gives the upper bound. □

7.2 Lower bounds.

Now suppose there exists $\gamma \geq \beta$, an observable φ and a positive function $C(\alpha)$ of α on a neighborhood U of 0 (so $C(\alpha) > 0$ if $\alpha \in U/\{0\}$) such that for all n, $\mu\big(\big|\frac{1}{n}S_n(\varphi) - \bar{\varphi}\big| > \alpha\big) \geq C(\alpha)n^{-\gamma}$. Our results in this section still assume the structure of a Young Tower and rely in part on the return time estimate $\nu_\Delta(R > n) \leq Cn^{-\beta}$ in order to truncate the tower at certain levels. As a consequence the parameter β appears in our expression for τ. We show if we take a window of length n^τ, $\tau < \frac{\beta}{\gamma+\beta\gamma+\beta}$ then the time-averaged fluctuation persists almost surely. In the case that γ limits to β then we require $\tau < \frac{1}{2+\beta}$. Comparing Theorem 7.1 and Theorem 7.2 there is a gap $\frac{1}{2+\beta} < \tau < \frac{1}{\beta}$ for which we don't know the almost sure limits or fluctuations of windows of length n^τ.

Theorem 7.2. *Suppose that (T, M, μ) is modeled by a Young Tower and $\nu_\Delta(R > n) \leq Cn^{-\beta}$ for some $\beta > 1$. Suppose that $\gamma \geq \beta$ and there exists a function C on a neighborhood U of $\alpha = 0$ such that $C(\alpha) > 0$ if $\alpha \in U/\{0\}$ and for all $n > 0$ and Lipschitz continuous function ϕ one has*

$$\mu\left(\left|\frac{1}{n}S_n(\varphi) - \bar{\varphi}\right| > |\alpha|\right) \geq C(\alpha)n^{-\gamma}$$

If $\alpha \in U$ and $0 < \tau < \frac{\beta}{\gamma+\gamma\beta+\beta}$ then for μ a.e. $x \in M$

$$\limsup_{n\to\infty} \max_{0\leq m\leq n-n^\tau} \left|n^{-\tau}S_{n^\tau} \circ T^m(x) - \bar{\phi}\right| \geq |\alpha|$$

Proof. We consider the case $\alpha > 0$ as the case $\alpha < 0$ is the same with obvious modifications. Let $0 < \epsilon \ll \alpha$ and put

$$A_{n^\tau}(\epsilon) = \left\{(x, j) \in \Delta : \sum_{r=1}^{n^\tau} \varphi \circ F^r(x, j) \leq n^\tau(\alpha - \epsilon)\right\}$$

If $(x', 0) \in W^s((x, 0))$ then $d(F^j(x, 0), F^j(x', 0)) \leq Kd(x, x')$ by Section 2 Assumptions (a) and (c), which implies that $|S_{n^\tau}\varphi(x, 0) - S_{n^\tau}(x', 0)| \leq n^\tau Kd(x, x')\|\varphi\|_{Lip}$. If $d(x, x') < \frac{\epsilon}{4K\|\varphi\|_{Lip}n^\tau}$ we obtain $|S_{n^\tau}\varphi(x, 0) - S_{n^\tau}(x', 0)| \leq \epsilon/4$. Consequently if $y = (x, j) \in A_{n^\tau}(\epsilon)$ and $d(x, x') < \frac{\epsilon}{4K\|\varphi\|_{Lip}n^\tau}$ then $y' = (x', j) \in A_{n^\tau}(3\epsilon/4)$.

We define $\rho_n := [\frac{1}{\ln \beta_1} \ln(\frac{\epsilon}{4K\|\varphi\|_{Lip}n^\tau})] + 1$, with this definition $\beta_1^{\rho_n} < \frac{\epsilon}{4K\|\varphi\|_{Lip}n^\tau}$.

Let

$$\delta > \frac{\tau\gamma}{\beta}$$

and define

$$B_{n^\tau}(\epsilon) = \left\{(x, 0) \in \Lambda : \exists\ 0 \leq j < n^\delta \text{ with } (f^{\rho_n}x, j) \in A_{n^\tau}(\epsilon)^c\right\} = f^{-\rho_n}(\Pi A_{n^\tau}(\epsilon)^c),$$

where $\Pi : \Delta \to \Lambda$ is the projection given by $\Pi((x, j)) = (x, 0)$ $(j < R(x))$. By the assumption of the theorem that

$$\mu\left(\left|\frac{1}{n}S_n(\varphi) - \bar{\varphi}\right| > \alpha\right) \geq C(\alpha)n^{-\gamma}$$

and the fact that $\mu = \pi^*\nu$ we have

$$\nu_\Delta(A_{n^\tau}(\epsilon)^c) \geq C(\alpha - \epsilon)n^{-\gamma\tau}.$$

We have also assumed that

$$\nu_\Delta(R > \ell) \leq C\ell^{-\beta}$$

and hence

$$\nu_\Delta(R > n^\delta) \le Cn^{-\beta\delta} = o(n^{-\delta})$$

as $\beta > 1$. Since $\nu_\Delta = \nu \times$ (counting measure) we get for $D \subset \Delta$

$$\nu(\Pi(D)) \ge \frac{\nu_\Delta(D) - \nu_\Delta(R > n^\delta)}{n^\delta}.$$

Consequently

$$\nu(\Pi(A_{n^\tau}(\epsilon)^c) \cap (R < n^\delta)) \ge \left(C(\alpha - \epsilon)n^{-\tau\gamma} - o(n^{-\delta\beta})\right) n^{-\delta}$$

and since $\delta\beta > \tau\gamma$ the first term dominates and we obtain

$$\nu(\Pi(A_{n^\tau}(\epsilon)^c \cap (R < n^\delta))) \ge C(\alpha - \epsilon)n^{-\tau\gamma-\delta}$$

Since f^{ρ_n} preserves ν,

$$\nu(f^{-\rho_n}\Pi(A_{n^\tau}(\epsilon)^c \cap (R < n^\delta)) \ge C(\alpha - \epsilon)n^{-\tau\gamma-\delta}.$$

Note that if $f^{\rho_n}(x, 0) \in \Pi(A_{n^\tau}(\epsilon)^c \cap (R < n^\delta))$ then there exists $j < n^\delta$ with $(f^{\rho_n}(x), j) \in A_{n^\tau}(\epsilon)^c$. Hence

$$\nu(B_{n^\tau}(\epsilon)) \ge C(\alpha - \epsilon)n^{-\tau\gamma-\delta}.$$

Note that

$$B_{n^\tau}(\epsilon)^c = \{(x, 0) \in \Lambda : \forall\ 0 \le j < n^\delta,\ (f^{\rho_n}x, j) \in A_{n^\tau}(\epsilon)\}$$

and define

$$\tilde{B}_{n^\tau}(\epsilon) = \bigcup_{x \in B_{n^\tau}(\epsilon)} W^s(x)$$

which by choice of n^δ implies that

$$\tilde{B}_{n^\tau}(\epsilon) \subset B_{n^\tau}(3\epsilon/4).$$

We will approximate $1_{\tilde{B}_{n^\tau}(\epsilon)^c}$ by a function $h_{n^\tau}(\epsilon)$ which has Lipschitz constant $\beta_1^{-n^{\tau'}}$, for $\tau' > \delta$, in the d_{β_1}-norm. For that purpose let $\tau' > \delta$ and define

$$h_{n^\tau}(\epsilon)(p) = \max(0, 1 - d((p, 0), \tilde{B}_{n^\tau}(\epsilon)^c)\beta_1^{-n^{\tau'}})$$

Assume $h_{n^\tau}(\epsilon)(p, 0) > 0$, then there exists $(q, 0) \in \tilde{B}_{n^\tau}(\epsilon)^c$ so that $d((p, 0), (q, 0)) \le \beta_1^{n^{\tau'}}$ which in particular implies that $s((p, 0), (q, 0)) \ge n^{\tau'}$ and allows us to conclude that $d(f^{\rho_n}(p, 0), f^{\rho_n}(q, 0)) \le \beta_1^{n^{\tau'}-\rho_n}$ as $\rho_n < n^\tau < n^{\tau'}$. Thus

$$d(F^j f^{n_1}(p, 0), F^j f^{n_1}(q, 0)) \le K\beta_1^{n^{\tau'}-\rho_n-j} \le K\beta_1^{n^{\tau'}-\rho_n-n^\delta}$$

by Section 2 Assumption (c). Consequently

$$|S_{n^\tau}\varphi(p, 0) - S_{n^\tau}\varphi(q, 0)| \le \|\varphi\|_{Lip} n^\tau K\beta_1^{n^{\tau'}-\rho_n-n^\delta} \le \frac{\epsilon}{4}$$

for all n large enough. Thus $(p, 0) \in B_{n^\tau}(\epsilon/2)^c$ which implies that the support of $h_{n^\tau}(\epsilon)$ is contained in $B_{n^\tau}(\epsilon/2)^c$.

Now we let $1 > \tau_1 > \tau' > \tau$ and consider

$$G_n(\epsilon) = \bigcap_{m=0}^{[n/n^{\tau_1}]} f^{-mn^{\tau_1}} B_{n^\tau}(\epsilon)^c$$

We will show that

$$\sum_n \nu(G_n(\epsilon)) < \infty$$

Now

$$\begin{aligned}
\nu(G_n(\epsilon)) &\leq \nu\left(\prod_{m=0}^{n^{1-\tau_1}} h_{n^\tau}(\epsilon) \circ f^{mn^{\tau_1}}\right) \\
&\leq \nu(h_{n^\tau}(\epsilon))\nu(G_{n-1}(\epsilon)) + c_3 \|h_{n^\tau}(\epsilon)\|_{\beta_1} |G_{n-1}(\epsilon)|_\infty \beta_3^{n^{\tau_1}} \\
&\leq [\nu(h_{n^\tau}(\epsilon))]^{n^{1-\tau_1}} + nC_3\beta_3^{n^{\tau_1}}\beta_1^{-n^{\tau'}}
\end{aligned}$$

The term $nC_3\beta_3^{n^{\tau_1}}\beta_1^{-n^{\tau'}}$ is summable in n as $\tau_1 > \tau' > \tau$. Using the fact that the support of $h_{n^\tau}(\epsilon)$ is contained in $B_{n^\tau}(\epsilon/2)^c$ the principal term is estimated by

$$[\nu(h_{n^\tau}(\epsilon))]^{n^{1-\tau_1}} \leq \Big(1 - C(\alpha - \frac{\epsilon}{2})n^{-\gamma\tau-\delta}\Big)^{n^{1-\tau_1}} \leq \exp\Big(-C(\alpha-\epsilon/2)n^{1-\tau_1-\gamma\tau-\delta}\Big),$$

as $\nu(h_{n^\tau}(\epsilon)) \leq \nu(B_{n^\tau}(\epsilon/2))$. Since τ_1 can be chosen arbitrarily close to τ and δ arbitrarily close to $\frac{\gamma\tau}{\beta}$ we may ensure that the power $1-\tau_1-\tau\gamma-\delta$ is positive for any chosen $\tau < \frac{\beta}{\gamma+\gamma\beta+\beta}$, which implies summability of $\nu(G_n(\epsilon))$.

Hence for ν a.e. $(x,0)$ there exists an $N(x)$ such that for all $n > N(x)$ there exists an $i < n - n^\tau$ with $f^i(x,0) \in B_{n^\tau}(\epsilon)$ and hence $(f^{\rho_n+i}x, j) \in A_{n^\tau}(\epsilon)^c$ for some $j < n^\delta$. Furthermore $i + j + \rho_n \leq n - n^{\tau_1} + n^\delta + C\ln n$.

Now we need to relate this conclusion, which holds for the map f, to the corresponding conclusion for the map $F : \Delta \to \Delta$. The maps f and F have different time clocks. Let $R_\ell = \sum_{i=0}^{\ell-1} R \circ f^i(x,0)$ denote the ℓ-th ergodic sum of R. Then $f^\ell(x,0) = F^{R_\ell}(x,0)$. We extend the definition of R_ℓ to the whole of Δ by defining $R_\ell(x,j) = [R(x,0) - j] + R_{\ell-1}f(x,0)$.

By Birkhoff's ergodic theorem

$$\lim_{n\to\infty} \frac{R_n(x,j)}{n} = \bar{R} = \frac{1}{\nu_\Delta(\Lambda)}$$

for ν_Δ a.e. $(x,j) \in \Delta$. We have shown that for ν_Δ a.e. (x,j)

$$\limsup \max_{1\leq j\leq N-\frac{N^\tau}{\bar{R}^\tau}} \frac{\bar{R}^\tau}{N^\tau} \sum_{i=j}^{j+\frac{N^\tau}{\bar{R}^\tau}} \varphi \circ F^j(x,0) \geq \alpha - \epsilon$$

which implies that for μ a.e. $p \in M$, for any $\tilde{\tau} < \tau < \frac{\beta}{\gamma+\gamma\beta+\beta}$

$$\limsup \max_{1\leq j\leq N-N^{\tilde{\tau}}} \frac{1}{N^{\tilde{\tau}}} \sum_{i=j}^{j+N^{\tilde{\tau}}} \varphi \circ T^j(p) \geq \alpha - \epsilon$$

and since $\tilde{\tau}, \tau$ were arbitrarily close to $\frac{\beta}{\gamma+\gamma\beta+\beta}$ this concludes the proof.

□

Example 7.3. *The condition* $\tau < \frac{\beta}{\gamma+\gamma\beta+\beta}$ *is close to optimal in that, taking* $\gamma = \beta$*, we require* $\tau < \frac{1}{2+\beta}$*. We may construct a Young Tower and observable* φ*,* $\int_\Delta \varphi \, d\nu_\Delta = 0$ *and* $\alpha > 0$ *such that* $\nu_\Delta(R > n) \leq Cn^{-\beta}$*, yet for all* $\tau > \frac{1}{\beta+1}$*,*

$$\lim_{n\to\infty} \max_{0\leq m\leq n-n^\tau} n^{-\tau} S_{n^\tau} \circ T^m = 0$$

We sketch the main idea of the tower and observable and make a couple of technical adjustments to ensure the tower is mixing and that the observable is not a coboundary. The construction is based on that of [1]. The base partition consists of disjoint intervals Λ_i *of length* $i^{-\beta-2}$*. Above the base element* Λ_i *the levels of the tower consist of* $\{(x,j) : 0 \leq j \leq 2i-1\}$*.*

Note that $\nu_\Delta(R > n) \leq \sum_{j=n/2}^{\infty}(2j)j^{-2-\beta} \leq Cn^{-\beta}$*. We define* φ *on the Tower by, if* $x \in \Lambda_i$*,*

$$\varphi(x,j) = \begin{cases} -1 & \text{if } 0 \leq j < i; \\ 1 & \text{if } i \leq j < 2i. \end{cases}$$

Clearly $\nu_\Delta(\varphi) = 0$*.*

Let $0 < \alpha < 1$*. Note that* $S_{n_\tau}\varphi(x,j) \geq n^\tau\alpha$ *only if* $(x,j) \in \{R > n^\tau\}$*.*

However if $\tau > \frac{1}{\beta+1}$ *then* $\sum_{n=1}^{\infty}\sum_{j\geq n^\tau} \bar{\nu}(\Lambda_j) \leq \sum_{n=1}^{\infty} n^{-\tau(\beta+1)} < \infty$*. Hence by the Borel-Cantelli lemma* $f^n(x,0) \in \bigcup_{j>n^\tau}\Lambda_j$ *only finitely many times for* $\bar{\nu}$ *a.e.* $(x,0)$*. This implies that for* $\bar{\nu}$ *a.e.* $(x,0)$ *there exists an* $N(x)$ *such that for all* $n \geq N(x)$

$$\text{for all } j < n : \sum_{r=0}^{n^\tau} \varphi(f^{j+r}x, 0) < \alpha n^\tau.$$

As $\alpha > 0$ *was arbitrary for* μ *a.e.* $x \in M$

$$\limsup_{n\to\infty} \max_{0\leq m\leq n-n^\tau} n^{-\tau} S_{n^\tau} \circ T^m = 0$$

The heights of the levels in the tower above are all multiples of 2 *so the tower is not mixing. Furthermore the observable* φ *is a coboundary and in fact if we define*

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

then one can check that

$$\varphi = \psi \circ F - \psi$$

But it is easy to adjust the tower so that the tower is mixing and φ *is not a coboundary, yet the pertinent features of the example remain.*

We will modify the tower and the observable so that the greatest common denominator of the return time function R *is* 1 *(to ensure the tower is mixing) and that the new observable is not a coboundary. We change* Λ_3 *to have height* 3*. This entails that the tower is mixing. On the levels above* Λ_3 *we modify* φ *to* φ_1 *so that* $\varphi_1(x,j) = \kappa > 0$*,* $j = 0,1,2$*,* $x \in \Lambda_3$ *where* $\kappa > 0$ *is small but* $\varphi_1 = \varphi$ *elsewhere . This entails* $r_1 := \nu_\Delta(\varphi_1) = \kappa\nu_\Delta(\Lambda_3) > 0$*. We subtract* $r_1/(\nu_\Delta(\Lambda_2))$ *from the value of* φ_1 *on* Λ_2 *to form a new observable* φ_2 *such that* $\nu_\Delta(\varphi_2) = 0$*. Since* F^3 *has a fixed point* p *on* Λ_3 *and since* $\sum_{j=0}^{2}\varphi_2(x,j) \neq 0$ *we conclude* φ_2 *is not a coboundary (by the Livšic theorem [19]). The new tower with observable* φ_2 *we defined has the properties of the former pertinent to our example.*

Acknowledgements We wish to thank an anonymous referee for a detailed, insightful reading and many helpful comments and suggestions.

References

[1] Bryc, W and Smolenski, W. On the Convergence of Averages of Mixing Sequences. Journal of Theoretical Probability, Vol 6, (1993), No. 3.

[2] Chen, H and Yu, M. A generalization of the Erdős-Rényi limit theorem and the corresponding multifractal analysis. J. Number Theory 192 (2018), 307-327.

[3] Chen, H, Ding, D and Long, X. The Hausdorff dimension of level sets described by Erdős-Rényi average. J. Math. Anal. Appl. 458 (2018), no. 1, 464-480.

[4] Chernov, N and Zhang, H. K. A family of chaotic billiards with variable mixing rates. Stoch. Dyn. 5 (2005) 535-553.

[5] Chazottes, J.-R.; Collet, P.: Almost-sure central limit theorems and the Erdős-Rényi law for expanding maps of the interval. Ergodic Theory Dynam. Systems **25**, (2005), 419–441.

[6] Czőrgő, S: Erdős-Rényi laws . Ann. Statistics. **7** (1979), 772–787.

[7] Deheuvels, P.; Devroye, L.; Lynch, J.: Exact convergence rate in the limit theorems of Erdős-Rényi and Shepp. Ann. Probab. **14** (1986), 209–223.

[8] Denker, M.; Kabluchko, Z.: An Erdős-Rényi law for mixing processes. Probab. Math. Statist. **27** (2007), 139–149.

[9] Denker, M and Nicol, M. Erdős-Rényi limit laws for hyperbolic dynamical systems. Journal LMS **87(2)** (2013), 497–508.

[10] Durrett, R. *Probability: Theory and Examples* (Cambridge Series in Statistical and Probabilistic Mathematics) 4th Edition. 2010.

[11] Erdős, P.; Rényi, A.: On a new law of large numbers. J. Analyse Math. **23** (1970), 103–111.

[12] Grigull, J.: *Große Abweichungen und Fluktuationen für Gleichgewichtsmaße rationaler Abbildungen*. Dissertation Georg-August-Universität zu Göttingen 1993.

[13] Kifer, Y. Functional Erdős-Rényi law of large numbers for nonconventional sums under weak dependence., Elect. J. Probab. **22** (2017), no.23, 1-17.

[14] Kifer,Y. Erdős-Rényi law of large numbers in the averaging setup., Stoch. Dyn. 18 (2018), no. 3, 1850018, 26 pp.

[15] A. Korepanov, Z. Kosloff and I. Melbourne. Explicit coupling argument for nonuniformly hyperbolic transformations. Proc. Roy. Soc. Edinburgh Sect. A 149 (2019), no. 1, 101-130.

[16] Melbourne, I.: Large and moderate deviations for slowly mixing dynamical systems. Proc. Amer. Math. Soc. **137** (2009), 1735–1741.

[17] Melbourne, I and Varandas, P. A note on statistical properties for non-uniformly hyperbolic systems with slow contraction and expansion. Stoch. Dyn. 16 (2016), no. 3, 1660012, 13 pp.

[18] Melbourne, I.; Nicol, M.: Large deviations for nonuniformly hyperbolic systems. Trans. Amer. Math. Soc. **360** (2008), 6661–6676.

[19] Nicol, M; Scott,A. Livšic theorems and stable ergodicity for group extensions of hyperbolic systems with discontinuities. Ergodic Theory Dynam. Systems 23 (2003), no. 6, 1867-1889.

[20] Pollicott, M; Sharp, R.: Large deviations for intermittent maps. Nonlinearity **22(9)** (2009), 2079–2090.

[21] Rey-Bellet, L.; Young, L.-S.: Large deviations in nonuniformly hyperbolic dynamical systems. Ergodic Theory Dynam. Systems **28** (2008) 587–612.

[22] Stein, E: *Singular integrals and differentiability properties of functions*, Princeton University Press, 1986.

[23] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998), 585–650.

[24] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999), 153–188.