

# MARTINGALE METHODS FOR MAXIMAL LARGE DEVIATIONS AND YOUNG TOWERS

JOSÉ F. ALVES, JOÃO S. MATIAS, AND IAN MELBOURNE

ABSTRACT. We develop a martingale approximation framework yielding quantitative maximal large deviations estimates for invertible dynamical systems. From suitable decay of correlations, we deduce these estimates and, as an application, we obtain Young structures with matching recurrence tails for partially hyperbolic diffeomorphisms with mostly expanding central direction. In a second application, we prove maximal large deviation estimates for systems modelled by Young towers with subexponential contraction and expansion. Many examples of slowly mixing billiards are covered by this result.

## 1. INTRODUCTION

One of the cornerstones of ergodic theory is *Birkhoff's Ergodic Theorem*, which ensures that if  $f: M \rightarrow M$  is an ergodic measure-preserving transformation on a probability space  $(M, \mu)$  and  $\varphi: M \rightarrow \mathbb{R}$  is integrable, then the time averages, or normalized Birkhoff sums, converge almost surely to the spatial averages:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi \circ f^j = \int \varphi d\mu.$$

This naturally raises the interesting question of how quickly this convergence takes place, that is, how fast the *large deviations* at time  $n$ ,

$$\text{LD}(\varphi, \varepsilon, n) = \mu \left( \left| \frac{1}{n} \sum_{j=0}^{n-1} \varphi \circ f^j - \int \varphi d\mu \right| > \varepsilon \right),$$

tend to zero as  $n \rightarrow \infty$ . Such estimates provide precise information about the “tails” of the distribution of Birkhoff sums and are closely related to the rate of decay of correlations and limit theorems.

---

2020 *Mathematics Subject Classification.* 37A05, 37A25, 37D25, 37D30, 60F05, 60F10.

*Key words and phrases.* Large Deviations, Maximal Large Deviations, Martingale Approximation, Decay of Correlations, Partial Hyperbolicity, Young Structures, Recurrence Rates.

JFA and JSM are partially supported by CMUP (UID/MAT/00144/2025) and PTDC/MAT-PUR/4048/2021, which are funded by FCT (Portugal) with national (MEC) and European structural funds through the program FEDER, under the partnership agreement PT2020. JFA is also supported by Royal Society Wolfson Visiting Fellowship RSWVF\25\R1\1006. JSM is also supported by the FCT doctoral scholarship 2021.07090.BD. The research of IM was supported in part by FAPESP grant number 2024/22093-5 at Instituto de Matemática e Estatística/IME/USP, São Paulo, and by the Sydney Mathematical Research Institute (SMRI-2026).

Regarding this speed of convergence, the classical problem was obtaining the *large deviation principle*, i.e., determining a rate function  $c(\varepsilon)$  such that

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log \text{LD}(\varphi, \varepsilon, n) = -c(\varepsilon).$$

Large deviation principles have been obtained for several classes of systems with hyperbolic behaviour. In the uniformly hyperbolic (Axiom A) setting, the theory is well understood, both for discrete- and continuous-time systems [27, 30, 37, 39]. For nonuniformly expanding/hyperbolic maps and flows, see for example [6, 7, 18, 19, 34].

More recently, several works have established a precise quantitative link between the decay of correlations and large deviations (LD). In particular, in [1, 2, 20, 32, 34] the authors showed that LD estimates can be derived from information on the decay of correlations, with comparable asymptotic rates. In the remainder of the introduction, we focus on polynomial rates for ease of exposition, but stretched exponential rates are covered equally from Section 2 onwards.

Given observables  $\varphi, \psi: M \rightarrow \mathbb{R}$ , their *correlation function* is defined by

$$\rho_{\varphi, \psi}(n) = \left| \int \varphi(\psi \circ f^n) d\mu - \int \varphi d\mu \int \psi d\mu \right|, \quad n \geq 1.$$

If there exists a constant  $\beta > 0$  such that

$$\rho_{\varphi, \psi}(n) \lesssim n^{-\beta},$$

then we speak of *polynomial decay of correlations*.<sup>1</sup> For noninvertible maps, the link between polynomial decay of correlations and polynomial LD was revealed in [32, 34]. An important consequence of quantitative LD estimates appeared in [2], where such bounds were used to construct (one-sided) Young towers with polynomial return-time tails. In this way, they obtained a converse to Young [41, Theorem 3].

Subsequently, using *maximal large deviations (MLD)*,

$$\text{MLD}(\varphi, \varepsilon, n) = \mu \left( \sup_{k \geq n} \left| \frac{1}{k} \sum_{j=0}^{k-1} \varphi \circ f^j - \int \varphi d\mu \right| > \varepsilon \right),$$

the authors in [10] were able to obtain sharper estimates for the return-time tails of the Young towers constructed in [2]. Clearly, the decay rate of  $\text{MLD}(\varphi, \varepsilon, n)$  implies a corresponding decay rate for  $\text{LD}(\varphi, \varepsilon, n)$ . As pointed out in [10], the converse automatically holds in the exponential and stretched exponential cases, but not in the polynomial case, leading to the improvement in the result of [2].

The results described above reveal a deep link for nonuniformly expanding maps between decay of correlations, MLD and geometric structures underlying the dynamics. One of our main goals in this work is to generalise such links to partially hyperbolic systems with contracting directions.

<sup>1</sup>Here  $a_n \lesssim b_n$  means that there exist constants  $C > 0$ ,  $n_0 \geq 1$  such that  $a_n \leq Cb_n$  for all  $n \geq n_0$ .

**Nonuniformly expanding maps.** We now describe more precisely the existing results in the noninvertible setting. Suppose that  $f$  is a piecewise smooth  $C^{1+\eta}$  map on a finite-dimensional Riemannian manifold  $M$  and that  $\mu$  is an  $f$ -invariant ergodic probability measure. Let  $\varphi: M \rightarrow \mathbb{R}$  be a fixed Hölder continuous observable.

- (1) By Young [41], if  $f$  is modelled by a (one-sided) Young tower with tails  $n^{-(\beta+1)}$ , then  $\rho_{\varphi,\psi}(n) \lesssim \|\psi\|_{\infty} n^{-\beta}$  for all  $\psi \in L^{\infty}(M)$ . This rate is optimal [24, 38].
- (2) By Melbourne & Nicol [32, 34], if  $\rho_{\varphi,\psi}(n) \lesssim \|\psi\|_{\infty} n^{-\beta}$  for all  $\psi \in L^{\infty}(M)$ , then  $\text{LD}(\varphi, \varepsilon, n) \lesssim n^{-\beta}$  and this rate is optimal.

Now suppose in addition that all Lyapunov exponents of  $f$  are positive. Then the circle of implications between Young tower, decay of correlations and LD can be completed.

- (3) By Alves *et al.* [2, 4], if  $\text{LD}(\varphi, \varepsilon, n) \lesssim n^{-\beta}$  for all Hölder observables  $\varphi$ , then  $f$  is modelled by a Young tower with tails  $n^{-(\beta-1)}$ .

As pointed out in Alves *et al.* [2], the combination of (2) and (3) implies that if correlations for Hölder observables against  $L^{\infty}$  observables decay at rate  $n^{-\beta}$ , then there exists a Young tower with polynomial tails  $n^{-(\beta-1)}$ . Moreover, they treated maps with critical points and singularities. In this way, they obtained a converse to the result of Young [41] in (1). Note however that this is not quite a perfect converse to (1) since there is a discrepancy of 2 in the polynomial degree. The situation was improved in [10] who proved

- (2') If  $\rho_{\varphi,\psi}(n) \lesssim \|\psi\|_{\infty} n^{-\beta}$  for all  $\psi \in L^{\infty}(M)$ , then  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-\beta}$ .
- (3') If  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-\beta}$  for all Hölder observables  $\varphi$ , then  $f$  is modelled by a Young tower with tails  $n^{-\beta}$ .

This still does not provide a perfect converse to (1), but the discrepancy in the polynomial degree has been reduced to 1. For polynomial return-time tails, this is currently the best result available.

**Nonuniformly hyperbolic diffeomorphisms.** Now suppose that  $f: M \rightarrow M$  is a  $C^{1+\eta}$  diffeomorphism on a finite-dimensional Riemannian manifold  $M$  and that  $\mu$  is an  $f$ -invariant ergodic probability measure. A longstanding problem has to been to obtain satisfactory generalisations of the noninvertible results above in the invertible setting. For background on Young towers [40, 41]; see Section A.

An analogue of (1) is due to Gouëzel [25] using ideas from [11], and is written down in [36, Appendix] and [29]: If  $f$  is modelled by a (two-sided) Young tower with tails  $n^{-(\beta+1)}$ , then  $\rho_{\varphi,\psi}(n) \lesssim \|\phi\|_{C^{\eta}} \|\psi\|_{C^{\eta}} n^{-\beta}$  for all Hölder observables  $\varphi, \psi$ . Unfortunately, it is well-known (see for example [35, Remark 6]) that it is no longer possible to obtain decay of correlations against all  $\psi \in L^{\infty}(M)$ . Moreover, decay of correlations against all Hölder observables  $\psi$  seems insufficient for proceeding further. A stronger analogue of (1) was obtained by Demers *et al.* [22]. Let  $L^{\infty}(\mathcal{F}_0)$  be the space of  $L^{\infty}$  observables that are constant along local stable leaves (the notation will be explained later on in Section 2).

- (1'') By [22], if  $f$  is modelled by a (two-sided) Young tower with tails  $n^{-(\beta+1)}$ , then  $\rho_{\varphi,\psi}(n)$  decays at rate  $n^{-\beta}$  for all  $\varphi$  Hölder and  $\psi \in L^{\infty}(\mathcal{F}_0)$ . Moreover, on average, stable leaves are contracted under  $\varphi \circ f^n$  at rate  $n^{-\beta}$ .

The first main result in this paper, Theorem A below, provides the desired analogue to (2'):

(2'') Let  $\varphi \in L^\infty(M)$ . Suppose that  $\rho_{\varphi,\psi}(n)$  decays at rate  $n^{-\beta}$  for all  $\psi \in L^\infty(\mathcal{F}_0)$  and on average, stable leaves are contracted under  $\varphi \circ f^n$  at rate  $n^{-\beta}$ . Then  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-\beta}$ .

Suppose in addition that  $\mu$  is an *SRB* measure supported on a partially hyperbolic set with exponentially contracting stable directions and with all Lyapunov exponents along  $E^{cu}$  positive.

(3'') By [5, Theorem A], if  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-\beta}$  for all Hölder observables  $\varphi$ , then  $f$  is modelled by a Young tower with tails  $n^{-\beta}$ .

It is worth noting that [3, Theorem D] plays a key role in the proof of [5, Theorem A]. In this way, for partially hyperbolic systems, we are able to obtain the desired converse (again up to the discrepancy of 1 in the polynomial degree) to (1''). This result, together with a version for stretched exponential rates is stated in Theorem B.

**MLD for systems with subexponential contraction and expansion.** The aforementioned results on polynomial LD/MLD hold in particular for noninvertible systems modelled by (one-sided) Young towers with polynomial tails. Again, an important long-standing question was to prove such results for invertible systems modelled by (two-sided) Young towers with polynomial tails. This is easily resolved when there is exponential contraction along stable leaves: any Hölder observable  $\varphi$  can be lifted to a sum  $\hat{\varphi} + b$  where  $\hat{\varphi}$  is dynamically Hölder on the tower and constant along stable leaves while  $b$  is an  $L^\infty$  coboundary [33, Lemma 3.2]. In this way, MLD for  $\varphi$  reduces to MLD for  $\hat{\varphi}$  which holds by the results for one-sided towers.

Unfortunately the assumption about exponential contraction along stable leaves is somewhat hidden in [33, Section 3, Assumption (A2)(i)], and hence [32, 34], causing some confusion (including to the authors of [32, 34] who incorrectly applied their results to certain billiard examples [34, Examples 1.4 and 1.5]). In general, LD/MLD for Hölder observables for invertible systems modelled by polynomial two-sided Young towers is previously unproved. This is resolved in our final main result, Theorem C. As a consequence, in Section 7, we obtain new results on LD/MLD for numerous examples of slowly mixing billiards (including the incorrectly-argued examples in [32, 34]).

The remainder of this paper is organised as follows. Our main results, Theorems A, B and C are stated precisely in Section 2. In Section 3, we prove a purely probabilistic result, Theorem 3.1, showing how to prove MLD via martingale approximations. Theorem A is then proved in Section 4. Theorem B is proved in Section 5 as an application of Theorem A together with [5, Theorem A]. Theorem C is proved in Section 6 as a second application of Theorem A. In Section 7, we treat a number of examples using Theorem C.

## 2. STATEMENT OF THE MAIN RESULTS

For the statement of our first main result, we suppose that  $f: M \rightarrow M$  is an invertible ergodic measure-preserving transformation defined on a probability space  $(M, \mu)$ . Let  $\mathcal{W}$  be a countable partition of  $M$  into measurable sets and let  $\mathcal{F}_0$  be the  $\sigma$ -algebra generated by  $\mathcal{W}$ .

Throughout this section, we consider two types of rate sequence  $r(n)$ :

**Definition 2.1. Stretched exponential rates:**  $r(n) = e^{-\tau n^\omega}$  where  $\tau > 0$  and  $0 < \omega \leq 1$  are constants; in this case we set  $r'(\varepsilon, n) = \exp\{-\tau' \varepsilon^\omega n^{\frac{\omega}{2}}\}$  for  $\tau' > 0$  chosen sufficiently small.

**Polynomial rates:**  $r(n) = n^{-\beta}$  for some constant  $\beta > 0$ ; in this case, fix any  $p > \max\{2, 2\beta\}$ , and define  $r'(\varepsilon, n) = \varepsilon^{-p} n^{-\beta}$ .

**Theorem A.** Assume that  $f^{-1}\mathcal{F}_0 \subset \mathcal{F}_0$  and let  $r, r'$  be as in Definition 2.1. Let  $\varphi: M \rightarrow \mathbb{R}$  be an  $L^\infty$  mean zero observable. Suppose that there is a constant  $C > 0$  such that

- (i)  $\rho_{\varphi, \psi}(n) \leq C \|\psi\|_\infty r(n)$ , for all  $\psi \in L^\infty(\mathcal{F}_0)$ ,  $n \geq 1$ ;
- (ii)  $\sum_{W \in f^n \mathcal{W}} \mu(W) \text{diam } \varphi(W) \leq Cr(n)$ , for all  $n \geq 1$ .

Then there exists a constant  $C' > 0$  such that

$$\text{MLD}(\varphi, \varepsilon, n) \leq C' r'(\varepsilon, n) \quad \text{for all } n \geq 1, \varepsilon > 0.$$

Our first application of Theorem A is in the partially hyperbolic setting with contracting directions. We give sufficient conditions based on rates of decay of correlations for the existence of a Young tower with specified tails.

Let  $f: M \rightarrow M$  be a diffeomorphism defined on a compact Riemannian manifold  $(M, d)$ . We say that a compact invariant set  $X \subset M$  is *partially hyperbolic set* if there exists a  $Df$ -invariant splitting

$$T_X M = E^s \oplus E^{cu}, \tag{2.1}$$

with  $\dim E^s \geq 1$ ,  $\dim E^{cu} \geq 1$ , together with constants  $C > 0$ ,  $0 < \lambda < 1$ , such that

- $E^s$  is *uniformly contracting*:  $\|Df|_{E_x^s}\| \leq C\lambda$ , for all  $x \in X$ ;
- $E^{cu}$  is *dominated* by  $E^s$ :  $\|Df|_{E_x^s}\| \cdot \|Df^{-1}|_{E_{f^{-1}x}^{cu}}\| \leq C\lambda$ , for all  $x \in X$ .

We refer to  $E^s$  as the *stable* subbundle and to  $E^{cu}$  as the *centre-unstable* subbundle. The stable bundle  $E^s$  integrates to a lamination  $\mathcal{W}^s$  consisting of stable leaves, and the  $\sigma$ -algebra  $\mathcal{F}_0$  generated by  $\mathcal{W}^s$  satisfies  $f^{-1}\mathcal{F}_0 \subset \mathcal{F}_0$ .

Recall that an observable  $\varphi: M \rightarrow \mathbb{R}$  is said to be  $\eta$ -Hölder continuous, for some  $\eta \in (0, 1]$ , if there exists a constant  $C > 0$  such that

$$|\varphi(x) - \varphi(y)| \leq Cd(x, y)^\eta, \quad \text{for all } x, y \in M.$$

Let  $C^\eta(M)$  denote the space of all  $\eta$ -Hölder continuous observables. The Hölder norm of an observable  $\varphi \in C^\eta$  is defined as

$$\|\varphi\|_{C^\eta} := \|\varphi\|_\infty + \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x, y)^\eta}.$$

The following theorem provides quantitative conditions under which the statistical properties of the system ensure the existence of a Young tower.

**Theorem B.** Let  $f: M \rightarrow M$  be a  $C^{1+\eta}$  diffeomorphism and let  $X \subset M$  be a partially hyperbolic set. Suppose that  $f$  admits an ergodic SRB measure  $\mu$  supported on  $X$  with all Lyapunov exponents along  $E^{cu}$  positive. Let  $r, r'$  be as in Definition 2.1.

Assume that there exists a constant  $C > 0$  such that

$$\rho_{\varphi, \psi}(n) \leq C \|\varphi\|_{C^\eta} \|\psi\|_\infty r(n), \quad \text{for all } \varphi \in C^\eta, \psi \in L^\infty(\mathcal{F}_0) \text{ and } n \geq 1.$$

Then  $f$  is modelled by a (two-sided) Young tower with

$$m_\gamma \{R > n\} \lesssim r'(1, n), \quad \text{for all } \gamma \in \Gamma^u.$$

Our second application of Theorem A is to obtain MLD for dynamical systems modelled by a Young tower with subexponential tails.

**Theorem C.** *Suppose that  $f: M \rightarrow M$  is an invertible dynamical system modelled by a (two-sided) Young tower with return time function  $R$ . Let  $\varphi \in C^\eta(M)$  and let  $\varepsilon > 0$ .*

- (a) *Assume that there exist constants  $\tau > 0$  and  $0 < \omega \leq 1$  such that  $m\{R > n\} \lesssim e^{-\tau n^\omega}$ . Let  $\omega' = \omega/(1+\omega)$ . Then  $\text{MLD}(\varphi, \varepsilon, n) \lesssim \exp\{-\tau' n^{\frac{\omega'}{2}}\}$  for some  $\tau' \in (0, \tau)$ .*
- (b) *If there exists  $\beta > 0$  such that  $m\{R > n\} \lesssim n^{-(\beta+1)}$ , then  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-\beta}$ .*

### 3. MLD VIA MARTINGALE APPROXIMATIONS

The goal of this section is to obtain LD/MLD for invertible maps via martingale approximations techniques in a purely probabilistic setting. Such martingale approximations were originally due to Gordin [23, 26]. As in [22], we use a refinement due to Dedecker *et al.* [21].

We consider both polynomial MLD and (stretched) exponential LD. (As already pointed out in [10], (stretched) exponential MLD is not a useful notion.)

Suppose that  $f: M \rightarrow M$  is an invertible measure-preserving transformation on the probability space  $(M, \mathcal{F}, \mu)$ .

**Theorem 3.1.** *Assume that  $\mathcal{F}_0 \subset \mathcal{F}$  is a  $\sigma$ -subalgebra satisfying  $f^{-1}\mathcal{F}_0 \subset \mathcal{F}_0$ . Let  $\varphi: M \rightarrow \mathbb{R}$  be an  $L^\infty$  mean zero observable.*

- (a) *Assume that there exist constants  $C > 0$ ,  $\tau > 0$  and  $0 < \omega \leq 1$  such that, for all  $p \geq 1$ ,  $n \geq 1$ ,*
  - (i)  $\|\mathbb{E}(\varphi \circ f^{-n} \mid \mathcal{F}_0)\|_p \leq C e^{-\frac{\tau}{p} n^\omega}$ ;
  - (ii)  $\|\mathbb{E}(\varphi \circ f^n \mid \mathcal{F}_0) - \varphi \circ f^n\|_p \leq C e^{-\frac{\tau}{p} n^\omega}$ .

*Then, there exists  $\tau' > 0$  such that*

$$\text{LD}(\varphi, \varepsilon, n) \leq 2 \exp\{-\tau' n^{\frac{\omega}{2}} \varepsilon^\omega\}, \quad \text{for all } n \geq 1, \varepsilon > 0.$$

- (b) *Assume that there exist constants  $C > 0$ ,  $\beta > 0$  and  $p > \max\{2, 2\beta\}$  such that, for all  $n \geq 1$ ,*
  - (i)  $\|\mathbb{E}(\varphi \circ f^{-n} \mid \mathcal{F}_0)\|_p \leq C n^{-\beta/p}$ ;
  - (ii)  $\|\mathbb{E}(\varphi \circ f^n \mid \mathcal{F}_0) - \varphi \circ f^n\|_p \leq C n^{-\beta/p}$ .

*Then, there exists  $K > 0$  such that*

$$\text{MLD}(\varphi, \varepsilon, n) \leq K \varepsilon^{-p} n^{-\beta}, \quad \text{for all } n \geq 1, \varepsilon > 0.$$

The proof of Theorem 3.1 is presented in the next two subsections.

**3.1. Stretched Exponential Case.** In this subsection, we establish Theorem 3.1(a). We assume throughout that  $\varphi: M \rightarrow \mathbb{R}$  is a mean zero  $L^\infty$  observable satisfying the hypotheses of Theorem 3.1(a). For  $n \geq 1$ , we define the Birkhoff sum

$$\varphi_n = \sum_{k=0}^{n-1} \varphi \circ f^k.$$

**Lemma 3.2.** *There exists a constant  $C > 0$  such that*

$$\left\| \max_{k \leq n} |\varphi_k| \right\|_p \leq Cp^{\frac{1}{\omega}} n^{\frac{1}{2}}, \quad \text{for all } n \geq 1 \text{ and } p > 2.$$

*Proof.* Let  $A_n = \sum_{j=1}^n \varphi \circ f^{-j}$ . Notice that for every  $k \leq n$ ,

$$\varphi_k = (A_n - A_{n-k}) \circ f^n.$$

Thus, we can deduce that

$$\left\| \max_{k \leq n} |\varphi_k| \right\|_p = \left\| \max_{k \leq n} |A_n - A_{n-k}| \right\|_p \leq 2 \left\| \max_{k \leq n} |A_k| \right\|_p, \quad (3.1)$$

which means that it suffices to estimate  $\left\| \max_{k \leq n} |A_k| \right\|_p$  to complete the proof.

In order to do so, we apply [21, Corollary 3.9]. (Note that  $T$  in [21] corresponds to our  $f^{-1}$ .) For  $r \geq 1$  and  $p > 2$ , this gives

$$\begin{aligned} \int \max_{k \leq 2^r} |A_k|^p d\mu &\leq 2^{\frac{rp}{2}} \int |\varphi|^p d\mu + 2^{\frac{rp}{2}} \left( \sum_{j=0}^{r-1} 2^{-\frac{j}{2}} \|\mathbb{E}(A_{2^j} | \mathcal{F}_0)\|_p \right)^p \\ &\quad + 2^{\frac{rp}{2}} \left( \sum_{j=1}^r 2^{-\frac{j}{2}} \|A_{2^j} - \mathbb{E}(A_{2^j} | f^{2^j} \mathcal{F}_0)\|_p \right)^p. \end{aligned} \quad (3.2)$$

We estimate each of the three terms individually. In what follows we use  $C_0$  to denote a positive constant whose value is independent of  $n$ ,  $p$  and  $r$ . The value of  $C_0$  may change from line to line.

For the first term, setting  $C_0 = \|\varphi\|_\infty 2^{-1/\omega}$ , we have

$$\|\varphi\|_p \leq \|\varphi\|_\infty = C_0 2^{1/\omega} \leq C_0 p^{1/\omega}. \quad (3.3)$$

For the second term in (3.2),

$$\begin{aligned} \sum_{j=0}^{r-1} 2^{-\frac{j}{2}} \|\mathbb{E}(A_{2^j} | \mathcal{F}_0)\|_p &\leq \sum_{j=0}^{r-1} 2^{-\frac{j}{2}} \sum_{k=1}^{2^j} \|\mathbb{E}(\varphi \circ f^{-k} | \mathcal{F}_0)\|_p \\ &= \sum_{k=1}^{2^{r-1}} \sum_{j=\lceil \log_2 k \rceil}^{r-1} 2^{-\frac{j}{2}} \|\mathbb{E}(\varphi \circ f^{-k} | \mathcal{F}_0)\|_p \\ &\leq (1 - 2^{-\frac{1}{2}})^{-1} \sum_{k=1}^{2^{r-1}} \|\mathbb{E}(\varphi \circ f^{-k} | \mathcal{F}_0)\|_p \leq C_0 \sum_{k=1}^{\infty} e^{-\frac{\tau}{p} k^\omega}, \end{aligned} \quad (3.4)$$

where in the final inequality we have used hypothesis (i) of Theorem 3.1(a). As in the proof of [1, Lemma 5], we deduce that

$$\sum_{k=1}^{\infty} e^{-\frac{\tau}{p} k^\omega} \leq \int_0^{\infty} e^{-\frac{\tau}{p} t^\omega} dt = \frac{1}{\omega} \left(\frac{p}{\tau}\right)^{\frac{1}{\omega}} \int_0^{\infty} e^{-s} s^{\frac{1}{\omega}-1} ds = C_0 p^{\frac{1}{\omega}}, \quad (3.5)$$

where we have performed the change of variables  $s = \frac{\tau}{p} t^\omega$ . Hence,

$$\sum_{j=0}^{r-1} 2^{-\frac{j}{2}} \|\mathbb{E}(A_{2^j} | \mathcal{F}_0)\|_p \leq C_0 p^{\frac{1}{\omega}}. \quad (3.6)$$

To estimate the third term in (3.2), we start by noticing that

$$\begin{aligned}
\|A_{2^j} - \mathbb{E}(A_{2^j} | f^{2^j} \mathcal{F}_0)\|_p &\leq \sum_{k=1}^{2^j} \|\varphi \circ f^{-k} - \mathbb{E}(\varphi \circ f^{-k} | f^{2^j} \mathcal{F}_0)\|_p \\
&= \sum_{k=1}^{2^j} \|\varphi \circ f^{2^j-k} - \mathbb{E}(\varphi \circ f^{2^j-k} | \mathcal{F}_0)\|_p \\
&= \sum_{k=0}^{2^j-1} \|\varphi \circ f^k - \mathbb{E}(\varphi \circ f^k | \mathcal{F}_0)\|_p.
\end{aligned} \tag{3.7}$$

Proceeding analogously to the arguments in (3.4) and (3.5), and using hypothesis (ii) of Theorem 3.1(a), we obtain

$$\sum_{j=1}^r 2^{-\frac{j}{2}} \|A_{2^j} - \mathbb{E}(A_{2^j} | f^{2^j} \mathcal{F}_0)\|_p \leq C_0 p^{\frac{1}{\omega}}. \tag{3.8}$$

Substituting (3.3), (3.6) and (3.8) into (3.2), we conclude that

$$\left\| \max_{k \leq 2^r} |A_k| \right\|_p \leq 2^{\frac{r}{2}} C_0 p^{\frac{1}{\omega}}.$$

Choosing  $r \geq 1$  such that  $2^{r-1} < n \leq 2^r$ , we deduce that

$$\left\| \max_{k \leq n} |A_k| \right\|_p \leq \left\| \max_{k \leq 2^r} |A_k| \right\|_p \leq 2^{\frac{r}{2}} C_0 p^{\frac{1}{\omega}} \leq C_0 p^{\frac{1}{\omega}} n^{\frac{1}{2}}.$$

By (3.1), this completes the proof.  $\square$

The next result extends the previous estimate to the range  $p \geq \omega$ .

**Corollary 3.3.** *There exists a constant  $C > 0$  such that*

$$\left\| \max_{k \leq n} |\varphi_k| \right\|_p \leq C p^{\frac{1}{\omega}} n^{\frac{1}{2}}, \quad \text{for all } n \geq 1 \text{ and } p \geq \omega.$$

*Proof.* The estimate holds for  $p \geq 3$  by Lemma 3.2. For  $\omega \leq p < 3$ , we have

$$\left\| \max_{k \leq n} |\varphi_k| \right\|_p \leq \left\| \max_{k \leq n} |\varphi_k| \right\|_3 \leq C 3^{\frac{1}{\omega}} n^{\frac{1}{2}} \leq (C 3^{\frac{1}{\omega}} \omega^{-\frac{1}{\omega}}) p^{\frac{1}{\omega}} n^{\frac{1}{2}}.$$

Hence the desired estimate holds for every  $p \geq \omega$ .  $\square$

**Lemma 3.4.** *There exists  $\tau' > 0$  such that  $\int \exp\{\tau' n^{-\frac{\omega}{2}} |\varphi_n|^\omega\} d\mu \leq 2$ , for all  $n \geq 1$ .*

*Proof.* By Corollary 3.3,

$$\begin{aligned}
\int \exp\{\tau' n^{-\frac{\omega}{2}} |\varphi_n|^\omega\} d\mu &= \sum_{j=0}^{\infty} \frac{(\tau')^j n^{-j\frac{\omega}{2}}}{j!} \int |\varphi_n|^{j\omega} d\mu \\
&\leq 1 + \sum_{j=1}^{\infty} \frac{(\tau')^j n^{-j\frac{\omega}{2}}}{j!} \left(C(j\omega)^{\frac{1}{\omega}} n^{\frac{1}{2}}\right)^{j\omega} = 1 + \sum_{j=1}^{\infty} (C^\omega \omega \tau')^j \frac{j^j}{j!}.
\end{aligned}$$

Since  $\left(1 + \frac{1}{j}\right)^j \leq e$ , by induction over  $j$ , we have  $j^j \leq e^j j!$  for every  $j \geq 1$ . Therefore,

$$\int \exp\{\tau' n^{-\frac{\omega}{2}} |\varphi_n|^\omega\} d\mu \leq 1 + \sum_{j=1}^{\infty} (C^\omega \omega e \tau')^j = 2$$

if we set  $\tau' = (2C^\omega \omega e)^{-1}$ .  $\square$

We are finally ready to complete the proof of Theorem 3.1(a).

*Proof of Theorem 3.1(a):* By Markov's inequality and Lemma 3.4,

$$\begin{aligned} \text{LD}(\varphi, \varepsilon, n) &= \mu\left(n^{-\frac{1}{2}} |\varphi_n| > n^{\frac{1}{2}} \varepsilon\right) = \mu\left(\exp\{\tau' n^{-\frac{\omega}{2}} |\varphi_n|^\omega\} > \exp\{\tau' n^{\frac{\omega}{2}} \varepsilon^\omega\}\right) \\ &\leq \exp\{-\tau' n^{\frac{\omega}{2}} \varepsilon^\omega\} \int \exp\{\tau' n^{-\frac{\omega}{2}} |\varphi_n|^\omega\} d\mu \leq 2 \exp\{-\tau' n^{\frac{\omega}{2}} \varepsilon^\omega\}, \end{aligned}$$

as required.  $\square$

**3.2. Polynomial Case.** In this subsection we prove Theorem 3.1(b). We assume throughout that  $\varphi: M \rightarrow \mathbb{R}$  is a mean zero  $L^\infty$  observable satisfying the hypotheses of Theorem 3.1(b).

We begin with a lemma very similar to Lemma 3.2.

**Lemma 3.5.** *There exists a constant  $C > 0$  such that*

$$\left\| \max_{k \leq n} |\varphi_k| \right\|_p \leq C n^{1-\frac{\beta}{p}}, \quad \text{for all } n \geq 1 \text{ and } p > \max\{2, 2\beta\}.$$

*Proof.* As in the (stretched) exponential case, let  $A_n = \sum_{j=1}^n \varphi \circ f^{-j}$ . Following the strategy of the previous section, let  $r \geq 1$ ,  $p > 2$ , so (3.2) holds. We estimate each term in the inequality (3.2) individually. Again,  $C_0$  denotes a positive constant whose value is independent of  $n$  and  $r$  and may change from line to line. However, this time  $C_0$  depends on  $p$ .

For the first term, simply take

$$\|\varphi\|_p \leq C_0. \quad (3.9)$$

For the second term, following the computations in (3.4),

$$\sum_{j=0}^{r-1} 2^{-\frac{j}{2}} \|\mathbb{E}(A_{2^j} | \mathcal{F}_0)\|_p \leq (1 - 2^{-\frac{1}{2}})^{-1} \sum_{k=1}^{2^{r-1}} k^{-\frac{1}{2}} \|\mathbb{E}(\varphi \circ f^{-k} | \mathcal{F}_0)\|_p \leq C_0 \sum_{k=1}^{2^r} k^{-\left(\frac{\beta}{p} + \frac{1}{2}\right)},$$

by hypothesis (i) of Theorem 3.1(b). Since  $p > \max\{2, 2\beta\}$ , we have  $\frac{\beta}{p} + \frac{1}{2} < 1$ , and so

$$\sum_{j=0}^{r-1} 2^{-\frac{j}{2}} \|\mathbb{E}(A_{2^j} | \mathcal{F}_0)\|_p \leq C_0 (2^r)^{\frac{1}{2} - \frac{\beta}{p}}. \quad (3.10)$$

Proceeding analogously and recalling (3.7), the third term in (3.2) is estimated by

$$\sum_{j=1}^r 2^{-\frac{j}{2}} \|A_{2^j} - \mathbb{E}(A_{2^j} | f^{2^j} \mathcal{F}_0)\|_p \leq C_0 (2^r)^{\frac{1}{2} - \frac{\beta}{p}}. \quad (3.11)$$

Substituting (3.9), (3.10) and (3.11) into (3.2), we conclude that

$$\left\| \max_{k \leq 2^r} |A_k| \right\|_p \leq 2^{\frac{r}{2}} C_0 (2^r)^{\frac{1}{2} - \frac{\beta}{p}} \leq C_0 (2^r)^{1 - \frac{\beta}{p}}.$$

Choosing  $r \geq 1$  such that  $2^{r-1} < n \leq 2^r$ , we deduce

$$\left\| \max_{k \leq n} |A_k| \right\|_p \leq \left\| \max_{k \leq 2^r} |A_k| \right\|_p \leq C_0 n^{1 - \frac{\beta}{p}}.$$

The result follows by (3.1).  $\square$

Define

$$M_n := \sup_{k \geq n} k^{-1} \left| \sum_{j=0}^{k-1} \varphi \circ f^j \right|, \quad n \geq 1.$$

We summarise the key argument in [10] for passing from large deviations to maximal large deviations as follows:

**Lemma 3.6.** *Let  $\phi \in L^\infty$ ,  $p \geq 1$ ,  $a > 0$ . Suppose there is a constant  $C > 0$  such that*

$$n^{-1} \left\| \max_{k \leq n} |\varphi_k| \right\|_p \leq C n^{-a}, \quad \text{for all } n \geq 1.$$

*Then there is a constant  $C' > 0$  such that*

$$\|M_n\|_p \leq C' n^{-a}, \quad \text{for all } n \geq 1.$$

*Proof.* The argument is almost identical to that presented in [10, pp. 8–9], so we just indicate the main steps. Define

$$C_n := 3 \frac{1}{3n} \left| \sum_{j=0}^{3n-1} \varphi \circ f^j \right| + 4 \frac{1}{2n} \max_{n \leq k \leq 2n-1} \left| \sum_{j=k}^{3n-1} \varphi \circ f^j \right|.$$

Then a calculation [10] yields  $M_n \leq M_{2n} + C_n$ . Since  $\lim_{n \rightarrow \infty} \|M_n\|_p = 0$  by the pointwise ergodic theorem and dominated convergence theorem (notice that  $\|M_n\|_\infty \leq \|\varphi\|_\infty$ ), it follows by induction<sup>2</sup> that

$$\|M_n\|_p \leq \sum_{j=0}^{\infty} \|C_{2^j n}\|_p.$$

Since  $\max_{k \leq n-1} \left| \sum_{i=k}^{n-1} \varphi \circ f^i \right| \leq 2 \max_{k \leq n} |\varphi_n|$ , it follows from Lemma 3.5 that

$$n^{-1} \left\| \max_{k \leq n-1} \left| \sum_{i=k}^{n-1} \varphi \circ f^i \right| \right\|_p \leq 2C n^{-a}, \quad \text{for all } n \geq 1.$$

Hence,  $\|C_n\|_p \lesssim n^{-a}$  and  $\|M_n\|_p \lesssim \sum_{j=0}^{\infty} (2^j)^{-a} n^{-a} \lesssim n^{-a}$ .  $\square$

*Proof of Theorem 3.1(b):* Let  $p > \max\{2, 2\beta\}$ . By Lemmas 3.5 and 3.6,  $\|M_n\|_p \lesssim n^{-\frac{\beta}{p}}$ . By Markov's inequality and Lemma 3.4,

$$\text{MLD}(\varphi, \varepsilon, n) = \mu(|M_n| > \varepsilon) \leq \varepsilon^{-p} \|M_n\|_p^p \lesssim \varepsilon^{-p} n^{-\beta},$$

as required.  $\square$

<sup>2</sup>The argument in [10] seems unnecessarily complicated here (at least for  $\varphi \in L^\infty$ ) and the induction works directly without introducing functions  $b_n$  and  $c_n$ .

## 4. DECAY OF CORRELATIONS AND MLD

In this section we prove Theorem A. First, we prove Theorem 4.1 below, which is a simple extension of [22, Theorem 3.1]. This result holds for both (stretched) exponential and polynomial rates of decay. For notational convenience, we denote the corresponding decay rate generically by  $r(n)$ ,  $n \geq 1$ .

We continue to suppose that  $f: M \rightarrow M$  is an invertible measure-preserving transformation on the probability space  $(M, \mu)$ . Suppose that  $\mathcal{W}$  is a countable partition of  $M$  into measurable sets and let  $\mathcal{F}_0$  be the  $\sigma$ -algebra generated by  $\mathcal{W}$ . We denote by  $L^\infty(\mathcal{F}_0)$  the space of essentially bounded  $\mathcal{F}_0$ -measurable functions.

**Theorem 4.1.** *Let  $\varphi: M \rightarrow \mathbb{R}$  be an  $L^\infty$  mean zero observable. Assume that there exists a constant  $C > 0$  and a decay rate  $r(n)$  such that*

- (i)  $\rho_{(\varphi, \psi)}(n) \leq C \|\psi\|_\infty r(n)$ , for all  $\psi \in L^\infty(\mathcal{F}_0)$ ,  $n \geq 1$ ;
- (ii)  $\sum_{W \in f^n \mathcal{W}} \mu(W) \text{diam } \varphi(W) \leq Cr(n)$ , for all  $n \geq 1$ .

Then, for all  $n \geq 1$  and  $p \geq 1$ ,

- (a)  $\|\mathbb{E}(\varphi \circ f^{-n} | \mathcal{F}_0)\|_p \leq C^{\frac{1}{p}} \|\varphi\|_\infty^{1-\frac{1}{p}} r(n)^{\frac{1}{p}}$ ;
- (b)  $\|\mathbb{E}(\varphi \circ f^n | \mathcal{F}_0) - \varphi \circ f^n\|_p \leq 2C^{\frac{1}{p}} \|\varphi\|_\infty^{1-\frac{1}{p}} r(n)^{\frac{1}{p}}$ .

*Proof.* Let  $\psi = |\mathbb{E}(\varphi \circ f^{-n} | \mathcal{F}_0)|^{p-1} \text{sgn}(\mathbb{E}(\varphi \circ f^{-n} | \mathcal{F}_0))$ . Then  $\psi \in L^\infty(\mathcal{F}_0)$  and  $\|\psi\|_\infty \leq \|\varphi\|_\infty^{p-1}$ . We recall the identities

$$\mathbb{E}(\varphi \circ f^{-n} | \mathcal{F}_0) \circ f^n = \mathbb{E}(\varphi | f^{-n} \mathcal{F}_0)$$

and thereby

$$\psi \circ f^n = |\mathbb{E}(\varphi | f^{-n} \mathcal{F}_0)|^{p-1} \text{sgn}(\mathbb{E}(\varphi | f^{-n} \mathcal{F}_0)).$$

Using these, we obtain

$$\begin{aligned} \|\mathbb{E}(\varphi \circ f^{-n} | \mathcal{F}_0)\|_p^p &= \|\mathbb{E}(\varphi | f^{-n} \mathcal{F}_0)\|_p^p = \int \{\mathbb{E}(\varphi | f^{-n} \mathcal{F}_0)\} (\psi \circ f^n) d\mu \\ &= \int \mathbb{E}\{\varphi(\psi \circ f^n) | f^{-n} \mathcal{F}_0\} d\mu = \int \varphi(\psi \circ f^n) d\mu \end{aligned}$$

Since  $\int \varphi d\mu = 0$ , we have shown that

$$\|\mathbb{E}(\varphi \circ f^{-n} | \mathcal{F}_0)\|_p^p = \rho_{\varphi, \psi}(n) \leq C \|\psi\|_\infty r(n) \leq C \|\varphi\|_\infty^{p-1} r(n),$$

which yields part (a).

Next, we recall by definition of conditional expectation that

$$\mathbb{E}(\varphi | f^n \mathcal{F}_0) = \sum_{W \in f^n \mathcal{W}} 1_W \mu(W)^{-1} \int_W \varphi d\mu.$$

Hence

$$\begin{aligned} |\mathbb{E}(\varphi|f^n\mathcal{F}_0) - \varphi|(x) &= \sum_{W \in f^n\mathcal{W}} 1_W(x) \left| \mu(W)^{-1} \int_W \varphi(y) d\mu(y) - \varphi(x) \right| \\ &\leq \sum_{W \in f^n\mathcal{W}} 1_W(x) \mu(W)^{-1} \int_W |\varphi(y) - \varphi(x)| d\mu(y) \\ &\leq \sum_{W \in f^n\mathcal{W}} 1_W(x) \text{diam } \phi(W). \end{aligned}$$

It follows that

$$\|\mathbb{E}(\varphi|f^n\mathcal{F}_0) - \varphi\|_1 \leq \sum_{W \in f^n\mathcal{W}} \mu(W) \text{diam } \phi(W) \leq Cr(n).$$

We deduce that

$$\begin{aligned} \|\mathbb{E}(\varphi \circ f^n | \mathcal{F}_0) - \varphi \circ f^n\|_p^p &= \|\mathbb{E}(\varphi|f^n\mathcal{F}_0) - \varphi\|_p^p \\ &\leq (2\|\varphi\|_\infty)^{p-1} \|\mathbb{E}(\varphi|f^n\mathcal{F}_0) - \varphi\|_1 \leq (2\|\varphi\|_\infty)^{p-1} Cr(n) \end{aligned}$$

which establishes part (b).  $\square$

*Proof of Theorem A:* The hypotheses of Theorem 3.1 follow from those of Theorem A by Theorem 4.1. Hence the result follows from Theorem 3.1.  $\square$

## 5. PARTIALLY HYPERBOLIC SYSTEMS WITH CONTRACTING DIRECTIONS

In this section, we apply Theorem A together with [5, Theorem A] to deduce the existence of a Young structure with specific return-time tails, thereby proving Theorem B.

We begin by proving a general result deducing maximal large deviations from decay of correlations for partially hyperbolic dynamical systems.

Let  $f: M \rightarrow M$  be a  $C^1$  diffeomorphism defined on a finite-dimensional Riemannian manifold  $(M, d)$  and let  $X \subset M$  be a compact invariant partially hyperbolic set with splitting (2.1). Let  $\mathcal{F}_0$  be the  $\sigma$ -algebra generated by  $\mathcal{W}^s$ .

**Proposition 5.1.** *Let  $\mu$  be an  $f$ -invariant ergodic probability measure on  $X$  and let  $r, r'$  be as in Definition 2.1. Let  $\varphi: M \rightarrow \mathbb{R}$  be a  $C^n$  mean zero observable. Assume that there exists a constant  $C > 0$  such that*

$$\rho_{\varphi, \psi}(n) \leq C \|\psi\|_\infty r(n), \quad \text{for all } \psi \in L^\infty(\mathcal{F}_0), n \geq 1.$$

*Then, there exists a constant  $C' > 0$  such that*

$$\text{MLD}(\varphi, \varepsilon, n) \leq C' r'(\varepsilon, n), \quad \text{for all } n \geq 1, \varepsilon > 0.$$

*Proof.* Let  $W^s(x)$  denote the stable leaf containing  $x \in X$ . Then  $fW^s(x) \subset W^s(fx)$  and it follows that  $f^{-1}W = \bigcup_{x \in f^{-1}W} W^s(x)$  for all  $W \in \mathcal{W}^s$ . Hence the condition  $f^{-1}\mathcal{F}_0 \subset \mathcal{F}_0$  is satisfied.

We are now in a position to apply Theorem A. Condition (i) in Theorem A is satisfied by assumption so it remains to verify condition (ii).

The stable lamination is exponentially contracting: there are constants  $C > 0, 0 < \lambda_1 < 1$  such that

$$d(f^n x, f^n y) \leq C \lambda_1^n d(x, y), \quad \text{for all } x, y \in X \text{ with } y \in W^s(x).$$

Let  $\lambda_2 = \lambda_1^\eta$ . Then

$$d(\varphi(f^n x), \varphi(f^n y)) \leq \|\varphi\|_{C^n} (C \operatorname{diam} M)^\eta \lambda_2^n, \quad \text{for all } x, y \in X \text{ with } y \in W^s(x).$$

That is,

$$\operatorname{diam} \varphi(f^n W) \leq \|\varphi\|_{C^n} (C \operatorname{diam} M)^\eta \lambda_2^n, \quad \text{for all } W \in \mathcal{W}^s.$$

In particular,  $\sum_{W \in f^n \mathcal{W}^s} \mu(W) \operatorname{diam} \varphi(W)$  decays exponentially, which certainly implies the subexponential contraction rate in condition (ii) in Theorem A. This completes the proof.  $\square$

*Proof of Theorem B.* We are in the situation of [5, Theorem A], from which it suffices to prove maximal large deviations rates as was done in Proposition 5.1. (In fact, to apply [5, Theorem A], we only require the estimates for  $\varepsilon = 1$ .)  $\square$

## 6. MLD FOR YOUNG TOWERS WITH SUBEXPONENTIAL EXPANSION/CONTRACTION

In this section, we prove Theorem C. Suppose that  $f: M \rightarrow M$  is an invertible dynamical system with a Young structure and return map  $F = f^R: Y \rightarrow Y$  as in Appendix A. Let  $(f_\Delta, \Delta, \mu_\Delta)$  and  $(\bar{f}_\Delta, \bar{\Delta}, \bar{\mu}_\Delta)$  be the associated two-sided and quotient Young towers. Recall the notions of dynamically Hölder observables on  $M$  and  $\Delta$  and  $d_\theta$ -Lipschitz observables on  $\bar{\Delta}$  from the Appendix.

**Theorem 6.1.** *Assume that  $\gcd\{R_i\} = 1$  so that  $\bar{\mu}_\Delta$  is mixing.*

(a) *Suppose that  $\bar{\mu}_\Delta(R > n) \lesssim e^{-\tau n^\omega}$  for some  $\tau > 0$ ,  $0 < \omega \leq 1$ . Then there exist  $\tau' \in (0, \tau)$  and  $C > 0$  such that*

$$\rho_{\bar{\varphi}, \bar{\psi}}(n) \leq C \|\bar{\varphi}\|_\theta \|\bar{\varphi}\|_\infty e^{-\tau' n^\omega} \quad \text{for all } d_\theta\text{-Lipschitz } \bar{\varphi}: \bar{\Delta} \rightarrow \mathbb{R}, \bar{\psi} \in L^\infty(\bar{\Delta}), n \geq 1.$$

(b) *Suppose that  $\bar{\mu}_\Delta(R > n) \lesssim n^{-(\beta+1)}$  for some  $\beta > 0$ . Then there exist  $C > 0$  such that*

$$\rho_{\bar{\varphi}, \bar{\psi}}(n) \leq C \|\bar{\varphi}\|_\theta \|\bar{\varphi}\|_\infty n^{-\beta} \quad \text{for all } d_\theta\text{-Lipschitz } \bar{\varphi}: \bar{\Delta} \rightarrow \mathbb{R}, \bar{\psi} \in L^\infty(\bar{\Delta}), n \geq 1.$$

*Proof.* See for example [29, Theorem 2.7(ii)] for part (a). Part (b) follows from [41] or [29, Theorem 2.7(i)].  $\square$

Let  $\mathcal{F}_0$  be the  $\sigma$ -algebra on  $\Delta$  generated by sets of the form  $(\bar{\pi}^{-1}E) \times \{\ell\}$  where  $E$  is a measurable subset of  $\bar{Y}_k$ ,  $k \geq 1$  and  $0 \leq \ell < R|\bar{Y}_k$ . Then  $f_\Delta^{-1}\mathcal{F}_0 \subset \mathcal{F}_0$ . Let  $L^\infty(\mathcal{F}_0)$  denote the space of observables  $\psi \in L^\infty(\Delta)$  such that  $\psi$  is  $\mathcal{F}_0$  measurable.

**Lemma 6.2.** *Fix a sequence  $r(n) > 0$ . Suppose that  $\rho_{\bar{\varphi}, \bar{\psi}}(n) \leq \|\bar{\varphi}\|_\theta \|\bar{\psi}\|_\infty r(n)$  for all  $d_\theta$ -Lipschitz  $\bar{\varphi}: \bar{\Delta} \rightarrow \mathbb{R}$  and all  $\bar{\psi} \in L^\infty(\bar{\Delta})$ . Then for any dynamically Hölder  $\varphi: \Delta \rightarrow \mathbb{R}$ , there exists  $C > 0$  such that*

$$\rho_{\varphi, \psi}(n) \leq C \|\psi\|_\infty r(n) \quad \text{for all } \psi \in L^\infty(\mathcal{F}_0) \text{ and } n \geq 1.$$

*Proof.* This is identical to the argument in [22, Section 3.2]. For completeness, we sketch the main steps.

Let  $L: L^1(\bar{\Delta}) \rightarrow L^1(\bar{\Delta})$  denote the transfer operator corresponding to  $\bar{f}: \bar{\Delta} \rightarrow \bar{\Delta}$ . It follows for instance from [29, Proposition 5.3] that we can choose a sequence of observables  $\varphi_\ell: \Delta \rightarrow \mathbb{R}$  such that

(i)  $\varphi_\ell$  is  $\mathcal{F}_0$ -measurable and hence projects to an observable  $\bar{\varphi}_\ell: \bar{\Delta} \rightarrow \mathbb{R}$ ;

- (ii)  $\sup_{\ell \geq 1} \|L^\ell \bar{\varphi}_\ell\|_\theta < \infty$ ;
- (iii)  $\lim_{\ell \rightarrow \infty} \|\varphi \circ f_\Delta^\ell - \varphi_\ell\|_1 = 0$ .

Let  $\psi \in L^\infty(\mathcal{F}_0)$  with projection  $\bar{\psi} \in L^\infty(\bar{\Delta})$ . Following [29, proof of Corollary 5.4],

$$\rho_{\varphi, \psi}(n) \leq |I_1(\ell, n)| + |I_2(\ell, n)| + I_3(\ell, n),$$

where

$$\begin{aligned} I_1(\ell, n) &= \int_{\Delta} (\varphi \circ f_\Delta^\ell - \varphi_\ell) \psi \circ f_\Delta^{\ell+n} d\mu_\Delta, \\ I_2(\ell, n) &= \int_{\Delta} (\varphi_\ell - \varphi \circ f_\Delta^\ell) d\mu_\Delta \int_{\Delta} \psi d\mu_\Delta, \quad I_3(\ell, n) = \rho_{\varphi_\ell, \psi}(\ell + n). \end{aligned}$$

By (iii),  $\lim_{\ell \rightarrow \infty} I_j(\ell, n) = 0$  uniformly in  $n$  for  $j = 1$  and  $j = 2$ . By (i) and the main hypothesis,

$$I_3(\ell, n) = \rho_{\bar{\varphi}_\ell, \bar{\psi}}(\ell + n) = \rho_{L^\ell \bar{\varphi}_\ell, \bar{\psi}}(n) \leq \|L^\ell \bar{\varphi}_\ell\|_\theta \|\bar{\psi}\|_\infty r(n) = \|L^\ell \bar{\varphi}_\ell\|_\theta \|\psi\|_\infty r(n).$$

By (ii),  $\sup_{\ell \geq 1} |I_3(\ell, n)| \lesssim r(n)$  and the result follows.  $\square$

**Corollary 6.3.** *Let  $\varphi : \Delta \rightarrow \mathbb{R}$  be dynamically Hölder.*

- (a) *Suppose that  $\mu_\Delta(R > n) \lesssim e^{-\tau n^\omega}$  for some  $\tau > 0$ ,  $0 < \omega \leq 1$ . Then there exist  $\tau' \in (0, \tau)$  and  $C > 0$  such that*

$$\rho_{\varphi, \psi}(n) \leq C \|\psi\|_\infty e^{-\tau' n^\omega} \quad \text{for all } \psi \in L^\infty(\mathcal{F}_0), n \geq 1.$$

- (b) *Suppose that  $\mu_\Delta(R > n) \lesssim n^{-(\beta+1)}$  for some  $\beta > 0$ . Then there exist  $C > 0$  such that*

$$\rho_{\varphi, \psi}(n) \leq C \|\psi\|_\infty n^{-\beta} \quad \text{for all } \psi \in L^\infty(\mathcal{F}_0), n \geq 1.$$

*Proof.* This is immediate from Theorem 6.1 and Lemma 6.2.  $\square$

*Proof of Theorem C.* Since  $\pi_\Delta : \Delta \rightarrow M$  is a measure-preserving semiconjugacy and Hölder observables  $\varphi : M \rightarrow \mathbb{R}$  lift to dynamically Hölder observables  $\varphi \circ \pi_\Delta : \Delta \rightarrow \mathbb{R}$ , it suffices to prove MLD for dynamically Hölder observables  $\varphi : \Delta \rightarrow \mathbb{R}$ .

We begin by proving the result under the additional assumption that  $\gcd\{R_i\} = 1$  (so  $\mu_\Delta$  and  $\bar{\mu}_\Delta$  are mixing). First, consider part (b) so  $\mu_\Delta(R > n) \lesssim n^{-(\beta+1)}$ . Let  $r(n) = n^{-\beta}$ . By Corollary 6.3(b), condition (i) of Theorem A is satisfied.

Define  $h_n(x) = \#\{0 \leq j \leq n : f_\Delta^j x \in Y\}$  for  $x \in \Delta$ . Then  $\text{diam } f_\Delta^n W^s(x) \lesssim \gamma^{h_n(x)}$  and hence  $\text{diam } \varphi(f_\Delta^n W^s(x)) \lesssim \gamma^{h_n(x)}$ . By [22, Lemma 3.2], condition (ii) of Theorem A is satisfied. Hence the result follows from Theorem A.

Next, consider part (a) so  $\mu_\Delta(R > n) \lesssim e^{-\tau n^\omega}$ . By Corollary 6.3, condition (i) of Theorem A is satisfied with  $r(n) = e^{-\tau' n^\omega}$ .

As in the proof of [22, Lemma 3.2],

$$\sum_{W \in f^n \mathcal{W}^s} \mu(W) \text{diam } \varphi(W) \lesssim \sum_{k=1}^{n+1} k \gamma^k \left( n \mu(R \geq n/k) + \int_Y 1_{\{R > n\}} R d\mu \right).$$

Now,  $\int_Y 1_{\{R>n\}} R d\mu = \sum_{j>n} j\mu(R=j) \lesssim e^{-\tau'n^\omega}$  for all  $\tau' \in (0, \tau)$ . Hence

$$\sum_{k=1}^{n+1} k\gamma^k \int_Y 1_{\{R>n\}} R d\mu \lesssim e^{-\tau'n^\omega}.$$

Next, choose  $\omega' \in (0, \omega)$  and write

$$\begin{aligned} \sum_{k=1}^{n+1} k\gamma^k \mu(R \geq n/k) &= \sum_{1 \leq k \leq n^{\omega'}} k\gamma^k \mu(R \geq n/k) + \sum_{n^{\omega'} < k \leq n} k\gamma^k \mu(R \geq n/k) \\ &\lesssim \mu(R \geq n^{1-\omega'}) + \sum_{k>n^{\omega'}} k\gamma^k \lesssim n e^{-\tau n^{(1-\omega')\omega}} + e^{-\tau'n^{\omega'}} \end{aligned}$$

for some  $\tau'$ . Taking  $\omega' = \omega/(1+\omega)$ , we obtain

$$\sum_{W \in f^n \mathcal{W}^s} \mu(W) \text{diam } \varphi(W) \lesssim e^{-\tau'n^{\omega/(1+\omega)}},$$

so condition (ii) of Theorem A is satisfied with  $r(n) = e^{-\tau'n^{\omega/(1+\omega)}}$ . Again, the result follows from Theorem A.

It remains to relax the assumption that  $\gcd\{R_i\} = 1$ . Suppose that  $\gcd\{R_i\} = s \geq 2$ . We decompose the tower into  $\Delta = \bigcup_{r=1}^s \Delta^{(r)}$  where each  $\Delta^{(r)}$  is an  $f_\Delta^s$ -invariant tower over  $Y$  with return time  $R^{(r)} = \frac{1}{s}R$ . Note that  $f_\Delta$  cyclicly permutes  $\Delta^{(1)}, \dots, \Delta^{(s)}$ . Moreover, we obtain mixing  $f_\Delta^s$ -invariant probability measures  $\mu_\Delta^{(r)} = s\mu_\Delta|_{\Delta^{(r)}}$  on  $\Delta^{(r)}$ .

Let  $\varphi_k = \sum_{j=0}^{k-1} \varphi \circ f_\Delta^j$ . For  $E \subset \Delta$  measurable,  $\mu_\Delta(E) = \frac{1}{s} \sum_{r=1}^s \mu_\Delta^{(r)}(E \cap \Delta^{(r)})$ . Hence,

$$\text{MLD}(\varphi, \varepsilon, n) = \frac{1}{s} \sum_{r=1}^s \mu_\Delta^{(r)} \left\{ x \in \Delta^{(r)} \sup_{k \geq n} \left| \frac{1}{k} \varphi_k(x) - \int_\Delta \varphi d\mu_\Delta \right| > \varepsilon \right\}. \quad (6.1)$$

Next, we note that  $\varphi_k = \sum_{j=0}^{\lfloor k/s \rfloor} \varphi_s \circ f_\Delta^{sj} + O(1)$  and  $\int_{\Delta^{(r)}} \varphi_s d\mu_\Delta^{(r)} = s \int_\Delta \varphi d\mu_\Delta$  for  $r = 1, \dots, s$ . It follows that on  $\Delta^{(r)}$ ,

$$\frac{1}{k} \varphi_k - \int_\Delta \varphi d\mu_\Delta = \frac{1}{s} \left\{ \frac{1}{\lfloor k/s \rfloor} \sum_{j=0}^{\lfloor k/s \rfloor - 1} \varphi_s \circ f_\Delta^{sj} - \int_{\Delta^{(r)}} \varphi_s d\mu_\Delta^{(r)} \right\} + O\left(\frac{1}{k}\right),$$

and hence

$$\sup_{k \geq n} \left| \frac{1}{k} \varphi_k - \int_\Delta \varphi d\mu_\Delta \right| = \frac{1}{s} \sup_{k \geq \lfloor n/s \rfloor} \left| \frac{1}{k} \sum_{j=0}^{k-1} \varphi_s \circ f_\Delta^{sj} - \int_{\Delta^{(r)}} \varphi_s d\mu_\Delta^{(r)} \right| + O\left(\frac{1}{n}\right). \quad (6.2)$$

By (6.1) and (6.2), MLD for  $\varphi$  on  $(f_\Delta, \Delta, \mu_\Delta)$  reduces to MLD for  $\varphi_s$  on  $(f_\Delta^s, \Delta^{(r)}, \mu_\Delta^{(r)})$ . Note that  $\varphi_s$  is dynamically Hölder whenever  $\varphi$  is dynamically Hölder. Since  $(f_\Delta^s, \Delta^{(r)}, \mu_\Delta^{(r)})$  is mixing, we have reduced to the mixing case  $\gcd\{R_i\} = 1$ .  $\square$

## 7. EXAMPLES: SLOWLY MIXING BILLIARDS

In this section, we use Theorem C to prove MLD for various slowly mixing billiard examples. For background material on billiards, we refer to [13]. The billiard domain, denoted by  $Q$ , is a compact connected subset of  $\mathbb{R}^2$  or  $\mathbb{T}^2$  with piecewise smooth boundary and the billiard flow  $f_t$  is defined on  $Q \times S^1$ . Fix a point  $q \in Q$  and a unit vector  $\psi \in S^1$ . Then  $q$  moves in straight lines with unit speed in direction  $\psi$  until reflecting (angle of reflection equalling the angle of incidence) off the boundary  $\partial Q$ . This defines a volume-preserving flow. A natural Poincaré section is given by  $M = \partial Q \times [-\pi/2, \pi/2]$  corresponding to collisions with  $\partial Q$  (with outgoing velocities in  $[-\pi/2, \pi/2]$ ). The Poincaré map  $f: M \rightarrow M$  is called the *collision map* or the *billiard map*. It preserves a probability measure  $\mu$ , equivalent to Lebesgue, called Liouville measure.

*Example 7.1* (Bunimovich stadia [9]). These are convex billiard domains  $Q \subset \mathbb{R}^2$  where  $\partial Q$  is a simple closed curve consisting of two semicircles  $C_1, C_2$  of radius 1 and two parallel line segments  $S_1, S_2$  of length  $L$  tangent to the semicircles.

By [31, 17], the billiard map  $f: M \rightarrow M$  is modelled by a Young tower with polynomial tails  $n^{-2}$ . By Theorem C, we obtain  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-1}$  for dynamically Hölder observables  $\varphi$ .

*Example 7.2* (Semi-dispersing billiards). These are billiard domains  $Q = R \setminus \Omega$  where  $R \subset \mathbb{R}^2$  is a rectangle and  $\Omega \subset \text{Int } R$  is a disjoint union of strictly convex regions with  $C^3$  boundaries and nonvanishing curvature.

By [15, 17], the billiard map  $f: M \rightarrow M$  is modelled by a Young tower with polynomial tails  $n^{-2}$ . By Theorem C, we obtain  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-1}$  for dynamically Hölder observables  $\varphi$ .

*Example 7.3* (Billiards with cusps and flat cusps). These are billiard domains  $Q \subset \mathbb{R}^2$  where  $\partial Q$  is a simple closed curve consisting of finitely many convex inwards  $C^3$  curves with nonvanishing curvature such that the interior angles at corner points are zero. By [14, 17], the billiard map  $f: M \rightarrow M$  is modelled by a Young tower with polynomial tails  $n^{-2}$ . By Theorem C, we obtain  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-1}$  for dynamically Hölder observables  $\varphi$ .

In [42], the nonvanishing curvature assumption was relaxed and it was shown that billiards with flat cusps are modelled by Young towers with polynomial tails  $n^{-(\beta+1)}$  where  $\beta \in (0, 1)$ . By Theorem C, we obtain  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-\beta}$  for dynamically Hölder observables  $\varphi$ .

*Example 7.4* (Bunimovich flowers). This is a class of billiards where  $\partial Q$  is a union of smooth curves that are either convex inwards (with bounded nonvanishing curvature) or arcs contained in semicircles, satisfying some additional technical assumptions, see [8, 15].

By [15, 28], the billiard map  $f: M \rightarrow M$  is modelled by a Young tower with polynomial tails  $n^{-3}$ . By Theorem C, we obtain  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-2}$  for dynamically Hölder observables  $\varphi$ .

*Example 7.5* (Dispersing billiards with flat points). Classical dispersing billiards have  $Q = \mathbb{T}^2 \setminus \Omega$  where  $\Omega$  is a disjoint union of strictly convex regions with  $C^3$  boundaries and nonvanishing curvature. By [40, 12], the billiard maps have exponential decay of correlations.

The nonvanishing curvature condition is relaxed in [16]: they allowed two flat points with a periodic orbit running between the flat points. The boundary near the flat points in [16] has profile  $\pm(1 + |x|^b)$  for some  $b > 2$ . By [16, 28], these billiard maps are modelled by Young towers with polynomial tails  $n^{-(\beta+1)}$  where  $\beta = (b + 2)/(b - 2) \in (1, \infty)$ . By Theorem C, we obtain  $\text{MLD}(\varphi, \varepsilon, n) \lesssim n^{-\beta}$  for dynamically Hölder observables  $\varphi$ .

## APPENDIX A. BACKGROUND MATERIAL ON YOUNG TOWERS

**A.1. Young structures.** Consider  $M$  to be a finite-dimensional compact Riemannian manifold, and  $f: M \rightarrow M$  a piecewise  $C^1$  diffeomorphism, possibly with critical, singular, or discontinuity points. Let  $d$  denote the distance on  $M$ , and let  $m$  represent the Lebesgue measure on the Borel subsets of  $M$ , both induced by the Riemannian metric. Given a submanifold  $\gamma \subset M$ , we denote by  $m_\gamma$  the Lebesgue measure on  $\gamma$ , induced by the restriction of the Riemannian metric to  $\gamma$ .

We say that  $\Gamma$  is a *continuous family* of  $C^1$  disks in  $M$  with  $\dim \Gamma = k$  if there are a compact metric space  $K$ , a unit disk  $D \subset \mathbb{R}^k$  for some  $1 \leq k \leq \dim M - 1$  and an injective continuous function  $\Phi: K \times D \rightarrow M$  such that

- $\Gamma = \{\Phi(\{x\} \times D) : x \in K\}$ ;
- $\Phi$  maps  $K \times D$  homeomorphically onto its image;
- $x \mapsto \Phi|_{\{x\} \times D}$  defines a continuous map from  $K$  into the space of  $C^1$  embeddings of  $D$  into  $M$ .

We say that a compact set  $Y \subset M$  has a *product structure* if there exist continuous families of  $C^1$  disks  $\Gamma^s$  (stable disks) and  $\Gamma^u$  (unstable disks), such that

- $Y = \left( \bigcup_{\gamma \in \Gamma^s} \gamma \right) \cap \left( \bigcup_{\gamma \in \Gamma^u} \gamma \right)$ ;
- $\dim \Gamma^s + \dim \Gamma^u = \dim M$ ;
- each  $\gamma \in \Gamma^s$  meets each  $\gamma \in \Gamma^u$  in exactly one point.

Let  $\gamma^*(y)$  denote the disk in  $\Gamma^*$  containing  $y \in Y$ , for  $* = s, u$ .

We say that  $Y_0 \subset Y$  is an *s-subset* if  $Y_0$  has a product structure with respect to families  $\Gamma_0^s$  and  $\Gamma_0^u$  such that  $\Gamma_0^s \subset \Gamma^s$  and  $\Gamma_0^u = \Gamma^u$ ; *u-subsets* are defined similarly.

The compact set  $Y$  has a *Young structure* or *two-sided Gibbs-Markov structure* if  $Y$  has a product structure such that

$$m_\gamma(Y \cap \gamma) > 0, \quad \text{for all } \gamma \in \Gamma^u,$$

and conditions **(Y<sub>1</sub>)**–**(Y<sub>5</sub>)** below are satisfied.

**(Y<sub>1</sub>) Markov:** there are pairwise disjoint *s*-subsets  $Y_1, Y_2, \dots \subset Y$  such that

- $m_\gamma((Y \setminus \bigcup_i Y_i) \cap \gamma) = 0$ , for all  $\gamma \in \Gamma^u$ ;
- for each  $i \geq 1$ , there is  $R_i \in \mathbb{N}$  such that  $f^{R_i}(Y_i)$  is a *u*-subset and, moreover, for all  $y \in Y_i$ ,

$$f^{R_i}(\gamma^s(y)) \subset \gamma^s(f^{R_i}y) \quad \text{and} \quad f^{R_i}(\gamma^u(y)) \supset \gamma^u(f^{R_i}y).$$

The *recurrence time*  $R: Y \rightarrow \mathbb{N}$  and *return map*  $F = f^R: Y \rightarrow Y$  are defined to be

$$R|_{Y_i} = R_i \quad \text{and} \quad F|_{Y_i} = f^{R_i}|_{Y_i}, \quad i \geq 1.$$

We remark that  $R$  and  $F$  are defined on a full  $m_\gamma$ -measure subset of  $Y \cap \gamma$ , for each  $\gamma \in \Gamma^u$ . Thus, there exists a set  $Y' \subset Y$  intersecting each  $\gamma \in \Gamma^u$  in a full  $m_\gamma$ -measure subset, such

that  $F^n y$  belongs to some  $Y_i$ , for all  $n \geq 0$  and  $y \in Y'$ . For points  $y, y' \in Y'$ , we define the *separation time*

$$s(y, y') = \min \{n \geq 0 : F^n y \text{ and } F^n y' \text{ lie in distinct } Y_i\text{'s}\}, \quad (\text{A.1})$$

with the convention that  $\min \emptyset = \infty$ . For definiteness, we set the separation time equal to zero for all other points.

For the remaining conditions, we consider constants  $C \geq 1$  and  $0 < \beta < 1$  depending only on  $f$  and  $Y$ .

- (Y<sub>2</sub>) **Contraction on stable disks:** for all  $i \geq 1$ ,  $\gamma \in \Gamma^s$  and  $y, y' \in \gamma \cap Y_i$ ,
- $d(F^n y, F^n y') \leq C\beta^n$ , for all  $n \geq 0$ ;
  - $d(f^j y, f^j y') \leq Cd(y, y')$ , for all  $1 \leq j \leq R_i$ .
- (Y<sub>3</sub>) **Expansion on unstable disks:** for all  $i \geq 1$ ,  $\gamma \in \Gamma^u$  and  $y, y' \in \gamma \cap Y_i$ ,
- $d(F^n y, F^n y') \leq C\beta^{s(y, y')-n}$ , for all  $n \geq 0$ ;
  - $d(f^j y, f^j y') \leq Cd(Fy, Fy')$ , for all  $1 \leq j \leq R_i$ .
- (Y<sub>4</sub>) **Bounded distortion:** for all  $i \geq 1$ ,  $\gamma \in \Gamma^u$  and  $y, y' \in \gamma \cap Y_i$ ,

$$\log \frac{\det DF|_{T_y \gamma}}{\det DF|_{T_{y'} \gamma}} \leq C\beta^{s(Fy, Fy')}.$$

Let  $\gamma, \gamma' \in \Gamma^u$ . The *stable holonomy map*  $\Theta_{\gamma, \gamma'} : \gamma \cap Y \rightarrow \gamma' \cap Y$  is given by

$$\Theta_{\gamma, \gamma'}(y) = \gamma^s(y) \cap \gamma', \quad y \in \gamma \cap Y.$$

- (Y<sub>5</sub>) **Regularity of the stable holonomy:** for all  $\gamma, \gamma' \in \Gamma^u$ , the measure  $(\Theta_{\gamma, \gamma'})_* m_\gamma$  is absolutely continuous with respect to  $m_{\gamma'}$  and its density  $\rho_{\gamma, \gamma'}$  satisfies

$$C^{-1} \leq \int_{\gamma' \cap Y} \rho_{\gamma, \gamma'} dm_{\gamma'} \leq C \quad \text{and} \quad \log \frac{\rho_{\gamma, \gamma'}(x)}{\rho_{\gamma, \gamma'}(y)} \leq C\beta^{s(y, y')},$$

for all  $y, y' \in \gamma' \cap Y$ .

Finally, we assume that  $\int_{Y \cap \gamma} R dm_\gamma < \infty$  for some, and hence all,  $\gamma \in \Gamma^u$ .

By Young [40, Proof of Theorem 1], there is an  $F$ -invariant Borel probability measure  $\mu_Y$  on  $Y$  with absolutely continuous conditional measures on unstable disks  $\gamma^u$ .

**A.2. Two-sided Young towers.** We define the *two-sided Young tower*  $\Delta = \{(y, \ell) : y \in Y, 0 \leq \ell < R(y)\}$  and tower map

$$f_\Delta : \Delta \rightarrow \Delta, \quad f_\Delta(y, \ell) = \begin{cases} (y, \ell + 1), & 0 \leq \ell < R(y) - 1 \\ (Fy, 0), & \ell = R(y) - 1 \end{cases},$$

with ergodic invariant probability measure  $\mu_\Delta = (\mu_Y \times \text{counting}) / \int_Y R d\mu_Y$ . Define the semiconjugacy

$$\pi_\Delta : \Delta \rightarrow M, \quad \pi_\Delta(y, \ell) = f^\ell y.$$

Then  $\mu = \pi_\Delta^* \mu_\Delta$  is an  $f$ -invariant probability measure on  $M$ . By Young [40, Theorem 1],  $\mu$  is an SRB measure:  $f$  has no zero Lyapunov exponents  $\mu$ -a.e. and  $\mu$  induces absolutely continuous conditional measures on unstable disks  $\gamma^u$ . In addition,  $\mu$  is mixing if and only if  $\gcd\{R_i\} = 1$ .

**A.3. Quotient towers.** We define an equivalence relation  $\sim$  on  $Y$  and on  $\Delta$  where  $(y, \ell) \sim (y', \ell)$  if  $y' \in \gamma^s(y)$ . Note that  $R : Y \rightarrow \mathbb{Z}^+$  is constant on equivalence classes. Let  $\bar{Y} = Y / \sim$ . Define the *quotient Young tower*  $\bar{\Delta} = \{(y, \ell) : y \in \bar{Y}, 0 \leq \ell < R(y)\} = \Delta / \sim$  and tower map

$$\bar{f}_\Delta : \bar{\Delta} \rightarrow \bar{\Delta}, \quad \bar{f}_\Delta(y, \ell) = \begin{cases} (y, \ell + 1), & 0 \leq \ell < R(y) - 1 \\ (\bar{F}y, 0), & \ell = R(y) - 1 \end{cases},$$

with ergodic invariant probability measure  $\bar{\mu}_\Delta = (\mu_{\bar{Y}} \times \text{counting}) / \int_{\bar{Y}} R d\mu_{\bar{Y}}$ . The natural projection  $\bar{\pi} : \Delta \rightarrow \bar{\Delta}$  defines a measure-preserving semiconjugacy. Again,  $\bar{\mu}$  is mixing if and only if  $\gcd\{R_i\} = 1$ .

**A.4. Observables.** The separation time  $s$  on  $\Delta$  is constant on equivalence classes and hence defines a separation time  $s$  on  $\bar{\Delta}$ . Then  $d_\theta(y, y') = \theta^{s(y, y')}$  is a metric on  $\bar{\Delta}$  for  $\theta \in (0, 1)$ . Given an observable  $\bar{\varphi} : \bar{\Delta} \rightarrow \mathbb{R}$ , define

$$\|\bar{\varphi}\|_\theta = \|\bar{\varphi}\|_\infty + \sup_{\substack{(y, \ell), (y', \ell) \in \bar{\Delta} \\ y \neq y'}} \frac{|\bar{\varphi}(y, \ell) - \bar{\varphi}(y', \ell)|}{d_\theta(y, y')}.$$

We say that  $\bar{\varphi}$  is *d<sub>θ</sub>-Lipschitz* if  $\|\bar{\varphi}\|_\theta < \infty$ .

Returning to the two-sided tower  $\Delta$ , we say that an observable  $\varphi : \Delta \rightarrow \mathbb{R}$  is *dynamically Hölder* if it is bounded and there exist constants  $C > 0$ ,  $\theta \in (0, 1)$  such that

$$|\varphi(y, \ell) - \varphi(y', \ell)| \leq C(d_Y(y, y') + d_\theta(\bar{\pi}y, \bar{\pi}y')) \quad (\text{A.2})$$

for all  $(y, \ell), (y', \ell) \in \Delta$ .

For the underlying system, an observable  $\varphi : M \rightarrow \mathbb{R}$  is *dynamically Hölder* if the lifted observable  $\varphi \circ \pi_\Delta : \Delta \rightarrow \mathbb{R}$  is dynamically Hölder. By definition of Young structure, Hölder observables on  $M$  are automatically dynamically Hölder (with  $\theta = \beta^\eta$  where  $\eta$  is the Hölder exponent).

## REFERENCES

- [1] AIMINO, R., AND FREITAS, J. M. Large deviations for dynamical systems with stretched exponential decay of correlations. *Port. Math.* 76, 2 (2019), 143–152.
- [2] ALVES, J. F., FREITAS, J. M., LUZZATTO, S., AND VAIENTI, S. From rates of mixing to recurrence times via large deviations. *Adv. Math.* 228, 2 (2011), 1203–1236.
- [3] ALVES, J. F., DIAS, C. L., LUZZATTO, S., AND PINHEIRO, V. SRB measures for partially hyperbolic systems whose central direction is weakly expanding. *J. Eur. Math. Soc. (JEMS)* 19 (2017), no. 10, 2911–2946.
- [4] ALVES, J. F., LUZZATTO, S. AND PINHEIRO, V. Markov structures and decay of correlations for non-uniformly expanding dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire* 22 (2005), 817–839.
- [5] ALVES, J. F., AND MATIAS, J. S. From decay of correlations to recurrence times in invertible dynamical systems. *arXiv preprint arXiv:2401.06024* (2024).
- [6] ARAÚJO, V. Semicontinuity of entropy, existence of equilibrium states and continuity of physical measures. *Discrete Contin. Dyn. Syst.* 17, 2 (2007), 371–386.
- [7] ARAÚJO, V., AND PACIFICO, M. J. Large deviations for non-uniformly expanding maps. *J. Stat. Phys.* 125, 2 (2006), 415–457.
- [8] BUNIMOVICH, L. A. The ergodic properties of billiards that are nearly scattering. *Dokl. Akad. Nauk SSSR* 211 (1973), 1024–1026.

- [9] BUNIMOVICH, L. A. On the ergodic properties of nowhere dispersing billiards. *Comm. Math. Phys.* 65 (1979), 295–312.
- [10] BUNIMOVICH, L. A., AND SU, Y. Maximal large deviations and slow recurrences in weakly chaotic systems. *Adv. Math.* 432 (2023), Paper No. 109267, 58.
- [11] CHAZOTTES, J. R. AND GOUËZEL, S. Optimal concentration inequalities for dynamical systems. *Comm. Math. Phys.* 316 (2012), 843–889.
- [12] CHERNOV, N. Decay of correlations and dispersing billiards. *J. Statist. Phys.* 94 (1999), 513–556.
- [13] CHERNOV, N. AND MARKARIAN, R. *Chaotic billiards*. Mathematical Surveys and Monographs **127**, American Mathematical Society, Providence, RI, 2006.
- [14] CHERNOV, N. AND MARKARIAN, R. Dispersing billiards with cusps: slow decay of correlations. *Comm. Math. Phys.* 270 (2007), 727–758.
- [15] CHERNOV, N. AND ZHANG, H.-K. Billiards with polynomial mixing rates. *Nonlinearity* 18 (2005), 1527–1553.
- [16] CHERNOV, N. AND ZHANG, H.-K. A family of chaotic billiards with variable mixing rates. *Stoch. Dyn.* 5 (2005), 535–553.
- [17] CHERNOV, N. AND ZHANG, H.-K. Improved estimates for correlations in billiards. *Comm. Math. Phys.* 77 (2008), 305–321.
- [18] CHUNG, Y. M., RIVERA-LETELIER, J. AND TAKAHASI, H. Large deviation principle in one-dimensional dynamics. *Invent. Math.* 218 (2019), 853–888.
- [19] COMMAN, H. AND RIVERA-LETELIER, J. Large deviation principles for non-uniformly hyperbolic rational maps. *Ergodic Theory Dynam. Systems* 31 (2011), 321–349.
- [20] CUNY, C., DEDECKER, J., AND MERLEVÈDE, F. Deviation and concentration inequalities for dynamical systems with subexponential decay of correlations. *Stoch. Dyn.* 23, 3 (2023), Paper No. 2350025, 18.
- [21] DEDECKER, J., MERLEVÈDE, F., AND PÈNE, F. Empirical central limit theorems for ergodic automorphisms of the torus. *ALEA Lat. Am. J. Probab. Math. Stat.* 10, 2 (2013), 731–766.
- [22] DEMERS, M., MELBOURNE, I., AND NICOL, M. Martingale approximations and anisotropic Banach spaces with an application to the time-one map of a Lorentz gas. *Nonlinearity* 33, 8 (2020), 4095–4113.
- [23] GORDIN, M. I. The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR* 188 (1969), 739–741.
- [24] GOUËZEL, S. Sharp polynomial estimates for the decay of correlations. *Israel J. Math.* 139 (2004), 29–65.
- [25] GOUËZEL, S. Private communication.
- [26] HALL, P. AND HEYDE, C. C. *Martingale Limit Theory and Its Application*. Probability and Mathematical Statistics. Academic Press, New York, 1980.
- [27] KIFER, Y. Large deviations in dynamical systems and stochastic processes. *Trans. Amer. Math. Soc.* 321, 2 (1990), 505–524.
- [28] KOMÁLOVICS, Á. AND BÁLINT, P. Improved estimates of statistical properties in some non-uniformly hyperbolic dynamical systems. *J. Math. Phys.* 66 (2025), Paper No. 122708, 19.
- [29] KOREPANOV, A., KOSLOFF, Z. AND MELBOURNE, I. Explicit coupling argument for nonuniformly hyperbolic transformations. *Proc. Roy. Soc. Edinburgh A* 149 (2019), 101–130.
- [30] LOPES, A. O. Entropy and large deviation. *Nonlinearity* 3, 2 (1990), 527–546.
- [31] MARKARIAN, R. Billiards with polynomial decay of correlations. *Ergodic Theory Dynam. Systems* 24 (2004), 177–197.
- [32] MELBOURNE, I. Large and moderate deviations for slowly mixing dynamical systems. *Proc. Amer. Math. Soc.* 137, 5 (2009), 1735–1741.
- [33] MELBOURNE, I., AND NICOL, M. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.* 260 (2005), 131–146.
- [34] MELBOURNE, I., AND NICOL, M. Large deviations for nonuniformly hyperbolic systems. *Trans. Amer. Math. Soc.* 360, 12 (2008), 6661–6676.
- [35] MELBOURNE, I., AND TÖRÖK, A. Central limit theorems and invariance principles for time-one maps of hyperbolic flows. *Comm. Math. Phys.* 229, 1 (2002), 57–71.

- [36] MELBOURNE, I. AND TERHESIU, D. Decay of correlations for nonuniformly expanding systems with general return times. *Ergodic Theory Dynam. Systems* 34 (2014), 893–918.
- [37] OREY, S., AND PELIKAN, S. Large deviation principles for stationary processes. *Ann. Probab.* 16, 4 (1988), 1481–1495.
- [38] SARIG, O. M. Subexponential decay of correlations. *Invent. Math.* 150 (2002), 629–653.
- [39] YOUNG, L.-S. Large deviations in dynamical systems. *Trans. Amer. Math. Soc.* 318, 2 (1990), 525–543.
- [40] YOUNG, L.-S. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math. (2)* 147, 3 (1998), 585–650.
- [41] YOUNG, L.-S. Recurrence times and rates of mixing. *Israel J. Math.* 110 (1999), 153–188.
- [42] ZHANG, H. K. Decay of correlations for billiards with flat points II: cusps effect. *Dynamical systems, ergodic theory, and probability: in memory of Kolya Chernov*, Contemp. Math. **698**, Amer. Math. Soc., Providence, RI, 2017, pp. 287–316.

JOSÉ F. ALVES, CENTRO DE MATEMÁTICA DA UNIVERSIDADE DO PORTO, RUA DO CAMPO ALEGRE 687, 4169-007 PORTO, PORTUGAL, AND DEPARTMENT OF MATHEMATICAL SCIENCES, LOUGHBOROUGH UNIVERSITY, LOUGHBOROUGH LE11 3TU, UNITED KINGDOM

*Email address:* jfalves@fc.up.pt

*URL:* <http://www.fc.up.pt/cmup/jfalves>

JOÃO S. MATIAS, CENTRO DE MATEMÁTICA DA UNIVERSIDADE DO PORTO, RUA DO CAMPO ALEGRE 687, 4169-007 PORTO, PORTUGAL.

*Email address:* up201504959@fc.up.pt

IAN MELBOURNE, MATHEMATICS INSTITUTE, UNIVERSITY OF WARWICK, COVENTRY CV4 7AL, UNITED KINGDOM.

*Email address:* I.Melbourne@warwick.ac.uk