DICHOTOMY RESULTS FOR EVENTUALLY ALWAYS HITTING TIME STATISTICS AND ALMOST SURE GROWTH OF EXTREMES

MARK HOLLAND, MAXIM KIRSEBOM, PHILIPP KUNDE, AND TOMAS PERSSON

ABSTRACT. Suppose $(f, \mathcal{X}, \mu)$ is a measure preserving dynamical system and $\phi\colon \mathcal{X} \to \mathbb{R}$ a measurable function. Consider the maximum process $M_n := \max\{X_1, \dots, X_n\}$, where $X_i = \phi \circ f^{i-1}$ is a time series of observations on the system. Suppose that (u_n) is a non-decreasing sequence of real numbers, such that $\mu(X_1 > u_n) \to 0$. For certain dynamical systems, we obtain a zero–one measure dichotomy for $\mu(M_n \leq u_n \text{ i.o.})$ depending on the sequence u_n. Specific examples are piecewise expanding interval maps including the Gauß map. For the broader class of non-uniformly hyperbolic dynamical systems, we make significant improvements on existing literature for characterising the sequences u_n. Our results on the permitted sequences u_n are commensurate with the optimal sequences (and series criteria) obtained by Klass (1985) for i.i.d. processes. Moreover, we also develop new series criteria on the permitted sequences in the case where the i.i.d. theory breaks down. Our analysis has strong connections to specific problems in eventual always hitting time statistics and extreme value theory.

1. INTRODUCTION

1.1. **General introduction and set up.** Consider a dynamical system $(\mathcal{X}, \mathcal{B}, \mu, f)$, where $(\mathcal{X}, \mathcal{B}, \mu)$ is a measure space equipped with a compatible metric which we denote by dist (that is, a metric such that open subsets of $\mathcal{X}$ are measurable), $f\colon \mathcal{X} \to \mathcal{X}$ is a measurable transformation, and μ is an f-invariant probability measure supported on $\mathcal{X}$. Given an observable $\phi\colon \mathcal{X} \to \mathbb{R}$, i.e. a measurable function, we consider the stationary stochastic process $X_1, X_2, \dots$ defined as

$$X_i = \phi \circ f^{i-1}, \quad i \geq 1,$$

and its associated maximum process M_n defined as

$$M_n = \max(X_1, \dots, X_n).$$

Extreme value theory is based on understanding the limiting behaviour of M_n, either almost surely or in distribution. We focus on the former

---

Date: November 1, 2023.

2010 *Mathematics Subject Classification.* 37E05, 37A50, 37D05, 60G70, 11J70.

Key words and phrases. Extreme value theory, eventually always hitting points, Robbins–Siegmund series criterion, extremal index.

This project was initiated at a "workshop on shrinking targets" at the University of Hamburg in November 2019. The workshop was funded by means of the University of Hamburg's status as a University of Excellence. We would like to thank the University for this financial support and opportunity. M. Holland acknowledges support from the EPSRC grant EP/P034489/1.

task of understanding almost sure growth rates for M_n. This is a form of strong law of large numbers for the maximum process (M_n). If μ is ergodic and ϕ is essentially bounded then almost surely, $M_n \to \operatorname{ess\,sup}\phi$ while if $\operatorname{ess\,sup}\phi = \infty$, $M_n \to \infty$ almost surely.

A fundamental problem is to determine optimal bounding sequences u_n and v_n such that almost surely there exists $N > 0$, with $v_n \leq M_n(x) \leq u_n$, for all $n \geq N$. (Here N depends on x). For independent, identically distributed (i.i.d.) random variables, this problem has been widely studied, e.g. [3, 14, 20, 38, 39]. The main difficultly is to find the lower bound sequence v_n. The upper bound sequence u_n is generally easier to establish from standard First and Second Borel–Cantelli Lemmas. Let us introduce some standard notations. For a sequence of sets (E_n), we define $(E_n \text{ i.o.})$ to be the set of points $x \in \mathcal{X}$ for which $x \in E_{n_k}$ for an infinite subsequence (n_k). Here 'i.o.' means *infinitely often*. We define $(E_n \text{ ev.})$ to be the set of points $x \in \mathcal{X}$ for which there exists $N > 0$ such that $x \in E_n$ for all $n > N$. Here 'ev.' means *eventually*. Now, for general non-decreasing sequences u_n the events $\{M_n > u_n \text{ i.o.}\}$ and $\{X_n > u_n \text{ i.o.}\}$ are equal (modulo a set of zero μ measure). Thus by the First Borel–Cantelli Lemma, if $\sum_n \mu(X_1 > u_n) < \infty$ we deduce that $\mu(M_n \leq u_n \text{ ev.}) = 1$. Moreover if a dynamical Borel–Cantelli property holds for (X_n), with $\sum_n \mu(X_1 > u_n) = \infty$ then $\mu(M_n \geq u_n \text{ i.o.}) = 1$.

1.2. **Dichotomy results for maxima.** For i.i.d. processes, a relevant criterion for a sequence (u_n) to be an eventual lower bound for M_n is given in particular by [39, Theorem 2], via the *Robbins–Siegmund series criterion.* This can be stated as follows. Suppose that $(\hat{X}_n)$ is an i.i.d. process, with probability measure P, and let u_n denote a non-decreasing sequence with $P(\hat{X}_1 > u_n) \to 0$, and $nP(\hat{X}_1 > u_n) \to \infty$. Then for the corresponding maximum process $\hat{M}_n$ we have the dichotomy

$$\sum_{n=1}^{\infty} P(\hat{X}_1 > u_n) e^{-nP(\hat{X}_1 > u_n)} < \infty \quad \Rightarrow \quad P(\hat{M}_n \geq u_n \text{ ev.}) = 1, \tag{1}$$

$$\sum_{n=1}^{\infty} P(\hat{X}_1 > u_n) e^{-nP(\hat{X}_1 > u_n)} = \infty \quad \Rightarrow \quad P(\hat{M}_n \geq u_n \text{ ev.}) = 0. \tag{2}$$

Moreover, when $P(\hat{X}_1 > u_n) \to c$, then $P(\hat{M}_n \leq u_n \text{ i.o.}) = 0$, while if

$$\liminf_{n \to \infty} nP(\hat{X}_1 > u_n) < \infty,$$

then $P(\hat{M}_n \leq u_n \text{ i.o.}) = 1$.

However, within a dynamical systems framework, and also for general dependent random variables, optimal bounds on almost sure growth rates of M_n are unknown in general. Recent progress on this problem in dynamical systems includes the works of [21, 25, 28] where dynamical Borel–Cantelli approaches are used to determine bounds on M_n for a wide class of dynamical systems, e.g. non-uniformly expanding maps, and hyperbolic systems. More recently, this problem has also been discussed indirectly in the analysis of eventual always hitting time statistics [22, 32, 33, 34, 36, 40]. For these latter papers, they consider a sequence of balls (B_n), and define an *eventually* *always hitting* (EAH) event $\mathcal{H}_{\mathrm{ea}}$ via

$$\mathcal{H}_{\mathrm{ea}} = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \bigcup_{k=0}^{m-1} f^{-k}(B_m). \tag{3}$$

Equivalently, $x \in \mathcal{H}_{\mathrm{ea}}$, if for the sequence $\mathbf{B} = (B_n)$, there exists $m_0(x) \in \mathbb{N}$, such that for all $m \geq m_0(x)$ we have

$$\{x, f(x), \ldots, f^{m-1}(x)\} \cap B_m \neq \emptyset.$$

The term eventually always hitting was coined by Kelmer in [32] where necessary and sufficient conditions for $\mathcal{H}_{\mathrm{ea}}$ to be of full measure are established in the context of discrete-time homogeneous flows on finite volume hyperbolic manifolds of constant negative curvature. Shortly afterwards Kelmer and Yu [34] extended the investigation to flows on higher-rank homogeneous spaces while Kelmer and Oh considered the case of geodesic flow on geometrically finite hyperbolic manifolds of infinite volume [33]. Also, Kleinbock and Wadleigh [41] studied the concept in the context of higher dimensional Diophantine approximations.

The problems addressed in [36, 40] include conditions placed on the sequence of measures $\mu(B_n)$ that lead to either $\mu(\mathcal{H}_{\mathrm{ea}}) = 0$ or $\mu(\mathcal{H}_{\mathrm{ea}}) = 1$. In fact, by ergodicity they show that $\mu(\mathcal{H}_{\mathrm{ea}})$ can only take these zero–one values. To link this directly to the maximum process (M_n), consider the observable $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$, where $\psi \colon [0, \infty) \to \overline{\mathbb{R}}$ is a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$. (For example, one can take $\psi(y) = -\log(y)$.) Then the event $\{X_1 > u_n\}$ corresponds directly to a target $B_n = B(\tilde{x}, r_n)$ with $u_n = \psi(r_n)$, and the event $\{M_n \leq u_n\}$ is the event $\cap_{k=1}^{n}\{X_k \leq u_n\}$. It follows that

$$\begin{aligned}\{M_n > u_n \text{ ev.}\} &:= \liminf_{n\to\infty} \left(\bigcap_{k=1}^{n} \{X_k \leq u_n\}\right)^{\complement} = \liminf_{n\to\infty} \bigcup_{k=1}^{n} \{X_k > u_n\} \\ &= \bigcup_{i=1}^{\infty} \bigcap_{n=i}^{\infty} \bigcup_{k=1}^{n} \{X_k > u_n\} = \mathcal{H}_{\mathrm{ea}}(\mathbf{B}).\end{aligned}$$

In this paper, we make several significant improvements on finding almost sure bounds for the maximum process M_n, and corresponding results for eventual always hitting time statistics via zero–one laws for the measure of $\mathcal{H}_{\mathrm{ea}}$. In particular we obtain dichotomy results consistent with the Robbins–Siegmund criteria described by Klass. Moreover we exhibit dynamical systems where the Robbins–Siegmund criteria are not valid, and we propose modified criteria beyond those stated in (1) and (2). We illustrate with a motivating example below. The main techniques we use are based upon ideas in *extreme value theory*, in particular on distributional convergence results for maxima, [12, 18, 27, 28, 45]. These methods generally differ to those used in obtaining dynamical Borel–Cantelli Lemmas alone.

In particular, we establish a dichotomy condition for $\mathcal{H}_{\mathrm{ea}}$ to be of full or zero measure in Theorem 3.2 for a class of interval maps. This is the first result on $\mathcal{H}_{\mathrm{ea}}$ with an exact dichotomy that we know of (see also [40, Question 7.1]).

Failure of Robbins–Siegmund series criterion. We broadly ascertain that conditions (1) and (2) are relevant to determine the almost sure growth bounds for maximum processes as generated from dynamical systems. However, we illustrate with a simple example to show that these conditions don't always apply. Let $(\hat{X}_n)$ be a i.i.d. process with continuous probability distribution function $F_{\hat{X}}(x) = 1 - 1/x$, with $x \in (0, \infty)$. For $n \geq 1$ define a new process (Y_n) by $Y_n = \max\{\hat{X}_n, \hat{X}_{n+1}\}$. The process (Y_n) is correlated only at short time lags, and indeed Y_n is independent of Y_m when $|n - m| \geq 2$. Conditions (1) and (2) apply to the process $(\hat{X}_n)$. By (1) we have for all $c < 1$,

$$\mu\Big(M_n^{\hat{X}} \geq \frac{cn}{\log\log n} \text{ ev.}\Big) = 1,$$

while for any $c' > 1$, (2) implies that

$$\mu\Big(M_n^{\hat{X}} \geq \frac{c'n}{\log\log n} \text{ ev.}\Big) = 0.$$

Here $M_n^{\hat{X}} = \max_{k \leq n} \hat{X}_k$. However, for the process (Y_n) we get corresponding statements for the maximum process $M_n^Y = \max_{k \leq n} Y_k$ by taking instead $c < 1/2$, and $c' > 1/2$. Such a result is inconsistent with conditions (1) and (2) when applied to the probability distribution for Y_n. This example is discussed more formally in Section 4.1.1. To gain insight into why conditions (1) and (2) fail for this example, we appeal to *extreme value theory* (EVT) surrounding existence of distributional limit laws for maxima. We next overview this topic.

1.3. **Background on distributional limit laws for extremes.** To obtain distributional limits in EVT, we seek sequences $a_n, b_n \in \mathbb{R}$ such that

$$\mu(\{\, x \in \mathcal{X} : a_n(M_n - b_n) \leq u \,\}) \to G(u),$$

for some non-degenerate distribution function $G(u)$, $-\infty < u < \infty$. Several results have shown that for sufficiently hyperbolic systems and for regular enough observables ϕ maximized at generic points $\tilde{x}$, the distribution limit is the same as that which would hold if $\{X_i\}$ were independent identically distributed (i.i.d.) random variables with the same distribution function as ϕ [16, 24, 28, 45]. Particular cases include laws towards Poisson type, described as follows. Suppose $\tau > 0$, and let $u_n(\tau)$ be a sequence such that

$$n\mu(X_1 > u_n(\tau)) \to \tau, \quad n \to \infty. \tag{4}$$

Then we say that an extreme value law with extremal index $\theta \in [0, 1]$ holds for M_n if

$$\mu(M_n \leq u_n(\tau)) \to e^{-\theta\tau}, \qquad n \to \infty. \tag{5}$$

If $(\hat{X}_n)$ is an i.i.d. process, then equation (5) holds for $\theta = 1$. Thus a non-trivial extremal index can only arise for dependent processes. Within EVT and wider statistical theory of extremes, the index measures the degree of *clustering* for a time series of maxima, see [43, 45] for details. Various methods are available to prove the convergence results above (in a dynamical systems context). An important method is a blocking algorithm approach, where in the context of general stationary stochastic processes see [14, 43]. For dynamical systems, a blocking method approach is described in [12].

To determine almost sure growth rates of maxima, we adapt the blocking method techniques that led to the distributional convergence results given by equation (5). As a naive approach, for a general sequence u_n, equations (4) and (5) lead us to compare $\mu(M_n \leq u_n)$ with $e^{-n\theta\mu(X_1>u_n)}$. In the i.i.d. case, we have the exact relation:

$$\mu(M_n \leq u_n) = (1 - \mu(X_1 > u_n))^n.$$

The right-hand side term is comparable to $e^{-n\mu(X_1>u_n)}$, assuming $n\mu(X_1 > u_n)^2 \to 0$. Thus, if we chose u_n so that the right-hand side is summable in n, then a First Borel–Cantelli Lemma implies $\mu(M_n \geq u_n \,\text{ev.}) = 1$. Thus, this relation is not so far from the first half of the Robbins–Siegmund criterion, namely (1). However, additional work is required to get the additional multiplier $\mu(X_1 > u_n)$ in (1). The second half of the criterion, namely (2) is much more delicate to obtain, even in the i.i.d. case. The issue being that $\{M_n \leq u_n\}$ is not a sequence of independent events, and hence a Second Borel–Cantelli Lemma cannot be readily applied to conclude whether or not $\mu(M_n \leq u_n \,\text{i.o.}) = 1$.

To obtain the relevant criteria (1) and (2) in the dynamical systems context, we also require a convergence rate in (5). This applies to the case $\theta = 1$ and also for case $\theta \neq 1$. We will treat these cases separately. Moreover, estimation of $\mu(M_n \leq u_n)$ is required for more general sequences u_n beyond those specified by equation (4).

Remark 1.1. Suppose $(\hat{X}_n)$ is an i.i.d. process with continuous distribution function $F_{\hat{X}}(x) = 1 - 1/x$, with $x \in (0, \infty)$. Then for the process $Y_n = \max\{\hat{X}_n, \hat{X}_{n+1}\}$ it can be shown that $\theta = 1/2$ in (5), see Section 4.1.1.

This remark suggests that examples for which the Robbins–Siegmund series criterion fails to apply are indeed those processes having a non-trivial extremal index $\theta \in (0, 1)$. This discussion is made more rigorous in Section 4 where we develop modified versions of (1) and (2) to account for processes having a non-trivial extremal index.

1.4. **Organisation of the paper and overview of results.** A complete theory is yet to be developed regarding dichotomy results on maxima and on eventually almost hitting time statistics. We now give an overview of the main results presented in this paper. In Section 2 we present a dichotomy result for piecewise expanding interval maps. This is Theorem 2.1, and the statement is consistent with criteria (1) and (2). As an application we consider the Gauß map, and obtain criteria applicable to determining the growth of the maximum for continued fraction expansion coefficients (associated to typical real numbers $x \in [0, 1]$).

In Section 3, we obtain dichotomy results for a broader class of interval maps, such as those having exponential decay of correlations in a suitable Banach space of functions. We show that dichotomy results of type (1) and (2) are applicable to a broad class of observable functions $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$ that attain their maximum at a generic point $\tilde{x} \in \mathcal{X}$. This is Theorem 3.2. Also within Section 3, we consider dynamical systems having weaker assumptions on the regularity of the invariant measure. For these systems, we obtain conditions close to the optimal sequences governed by

(1) and (2). For example, we show that $\mu(M_n \leq u_n \,\text{i.o.}) = 0$, provided u_n satisfies $\mu(X_1 > u_n) > c \log\log n/n$ for some $c > 1$, see Theorem 3.4. This gives improvements relative to the methods derived from dynamical Borel–Cantelli Lemma analysis, such as in [29, 36], where they require u_n to satisfy conditions of the form $\mu(X_1 > u_n) > (\log n)^\beta/n$, for some $\beta > 2$.

In Section 4, we obtain results that go beyond what we expect to see for i.i.d. processes. For the systems we consider we propose and apply modified criteria relative to (1) and (2). Such criteria incorporate an extremal index θ. See Theorem 4.1 for a precise statement. For the dynamical systems and observables we consider, the mechanisms leading to a non-trivial extremal index are driven by periodic behaviour. We show that our conditions are applicable to other dependent stochastic processes, where the extremal index is created due to other (clustering) mechanisms. We conjecture that our conditions are more widely applicable to other dependent processes.

In Section 5, we discuss higher dimensional dynamical systems such as those modelled by Young towers, [50]. Again, we obtain criteria towards (1). Relative to interval maps, we also need to consider regularity of the ergodic invariant measure. In general this measure can be singular with respect to Lebesgue measure. This creates obstacles when trying to develop and apply a version of e.g. (2). We obtain partial results, see Theorems 5.1 and 5.2.

Section 6 and onwards we devote to the proofs. In particular for Sections 6 and 7 we overview the proof strategy, including an overview of the blocking argument, such as the one developed in [12]. In the later sections, such as Section 13 we show that the dynamical assumptions stated in the main theorems are satisfied for a broad class of systems.

2. A dichotomy result for piecewise expanding maps

Our aim is to recover versions of the Robbins–Siegmund series critera (1) and (2), as applied to the maximum process $M_n = \max_{k \leq n-1} \phi(f^k)$, where we consider a measure preserving system $(f, \mathcal{X}, \mu)$, and μ is an ergodic measure. In this section, we will present our results for piecewise expanding interval maps.

Theorem 2.1. *Suppose that $f \colon \mathcal{X} \to \mathcal{X}$ is a piecewise expanding interval map with an ergodic measure μ which is absolutely continuous with respect to Lebesgue measure.*

Consider the observable $\phi(x) = \psi(\operatorname{dist}(x, \tilde{x}))$, where $\psi \colon [0, \infty) \to \overline{\mathbb{R}}$ is a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$, and a non-increasing sequence (r_n) such that $n \mapsto n\mu(B(\tilde{x}, r_n))$ is non-decreasing, $r_n = O(n^{-\sigma})$ for some $\sigma > \frac{4}{5}$, and such that for any $t > 0$ we have

$$\limsup_{k\to\infty} \frac{r_{\lfloor k^t \rfloor}}{r_{\lfloor (k+1)^t \rfloor}} < \infty. \tag{6}$$

Then, for μ-a.e. $\tilde{x}$ we have the following dichotomy:

(1) If the sequence (r_n) satisfies

$$\sum_{n=1}^{\infty} \mu(B(\tilde{x}, r_n)) e^{-n\mu(B(\tilde{x}, r_n))} < \infty,$$

then
$$\mu(\mathcal{H}_{\rm ea}) = \mu\big(\psi(r_n) \leq M_n \; ev.\big) = 1.$$

(2) If the sequence (r_n) *satisfies*
$$\sum_{n=1}^{\infty} \mu(B(\tilde{x}, r_n)) e^{-n\mu(B(\tilde{x},r_n))} = \infty,$$
then
$$\mu(\mathcal{H}_{\rm ea}) = \mu\big(\psi(r_n) \leq M_n \; ev.\big) = 0.$$

Remark 2.2. Since $n \mapsto n\mu(B(\tilde{x}, r_n))$ is non-decreasing it follows that $\sum_{n=1}^{\infty} \mu(B(\tilde{x}, r_n)) = \infty$. Hence in case (2) we also have
$$\mu\big(\psi(r_n) \leq M_n \text{ i.o.}\big) = \mu\big(\psi(r_n) \geq M_n \text{ i.o.}\big) = 1$$
because a dynamical Borel–Cantelli lemma holds in this setting [35].

Remark 2.3. As stated, the set of points $\tilde{x} \in \mathcal{X}$ for which (1) and (2) in Theorem 2.1 hold has full μ-measure. In general, it is not straightforward to know whether a particular $\tilde{x}$ is in this set. As we discuss in Section 4 periodic points are not in this full measure set. For such periodic points alternative formulations of (1) and (2) are required.

Theorem 2.1 is a consequence of Theorem 3.2 in the next section, combined with Proposition 13.4. We remark that the condition (6) is quite a mild condition on the sequence r_n. If r_n is regularly varying, (in the sense of [6]), then it will satisfy (6). Certain sequences with fast decay (such as exponential) violate (6), but this becomes a moot issue since we assume $\sum_n \mu(B(\tilde{x}, r_n)) = \infty$. Therefore r_n cannot decay too quickly, unless the measure density is quite degenerate at $\tilde{x}$. For i.i.d. processes, and depending on the criteria being used, mild regularity constraints are also discussed (and imposed) in [3, 38, 39]. The latter reference gives the most freedom on the allowed sequences, as we have already summarised in Section 1.

In the next section we present several results of this type which hold under various more or less abstract assumptions on the systems. For piecewise expanding systems, Proposition 13.4 tells us that these assumptions are satisfied for a.e. $\tilde{x}$. The restriction $\sigma > 4/5$ is a consequence of the methods of proof. We have not tried to optimise this range on σ. For completeness, upper bounds on the growth of maxima can be deduced from the following theorem.

Theorem 2.4. *Suppose that* $f \colon \mathcal{X} \to \mathcal{X}$ *is a dynamical system with an ergodic probability measure* μ. *If* (r_n) *is a sequence such that*
$$\sum_{n=1}^{\infty} \mu(B(\tilde{x}, r_n)) < \infty,$$
then
$$\mu\big(\psi(r_n) \geq M_n \; ev.\big) = 1.$$

Proof. Since $\mu(B(\tilde{x}, r_n))$ is summable, we get by the First Borel–Cantelli Lemma that almost surely, the event $\{X_n \geq \psi(r_n)\}$ happens only finitely many times. It follows that almost surely, $M_n \leq \psi(r_n)$ holds for all large enough n. □

The Gauß Map and growth of continued fractions. We end this section with an important application of Theorem 2.1, namely to the Gauß map. This allows us to obtain a dichotomy result on the almost sure growth rates of continued fraction expansion coefficients. Recall that for a number $x \in [0,1]$, its continued fraction is given by $x = [a_0, a_1, a_2, \ldots]$, where

$$a_k(x) = \left\lfloor \frac{1}{G^k(x)} \right\rfloor, \qquad \text{and} \qquad G(x) = \frac{1}{x} \mod 1.$$

The map $G\colon [0,1] \to [0,1]$ (with $G(0) = 0$) is the *Gauß map.* The map is piecewise expanding, full branch, with countable Markov partition. The map admits an ergodic measure μ with invariant density $\rho(x) = (\log 2)^{-1}(1+x)^{-1}$. In Philipp [47, Theorem 1] it is shown that for μ-almost all x

$$\liminf_{n\to\infty} n^{-1} L_n(x) \log\log n = \frac{1}{\log 2}, \tag{7}$$

where $L_n(x) = \max_{i\leq n} a_i(x)$. In Corollary 2.5 below we obtain a result commensurate with that of Philipp's dichotomy result [47, Theorem 2], which allows us to obtain higher order terms in the convergence rate to the limit. Philipp's dichotomy result naturally builds upon the earlier works of Barndorff-Neilson [3] as applied in the i.i.d. case. In the recent work of [40] they also obtain estimates which lead to the result of (7), but they don't obtain a sharp dichotomy criterion.

Corollary 2.5. *Suppose that $G\colon [0,1] \to [0,1]$ is the Gauß map, and μ the ergodic absolutely continuous invariant measure. Consider the observable $\phi(x) = \psi(\mathrm{dist}(x,\tilde{x}))$, where $\psi\colon [0,\infty) \to \overline{\mathbb{R}}$ is a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$, and a non-increasing sequence (r_n) such that $n \mapsto n\mu(B(\tilde{x}, r_n))$ is non-decreasing, $r_n = O(n^{-\sigma})$ for some $\sigma > \frac{4}{5}$, and satisfying (6). Then for μ-a.e. $\tilde{x}$ cases (1) and (2) of Theorem 2.1 apply.*

To relate this corollary to continued fractions, we take the observable $\psi(x) = \lfloor 1/x \rfloor$ with $\tilde{x} = 0$ so that $a_k(x) = \psi(G^k(x))$. One then attempts to apply Theorem 2.1, but in order to do so we need to know that we can use the theorem for $\tilde{x} = 0$. Instead we use Theorem 3.2, and we need only to check that condition (A2) (see Section 3) is satisfied for $\tilde{x} = 0$. To do so is standard, and is left out. See also [23]. The results of Philipp are recovered from Corollary 2.5 by first noting that

$$\mu([0, r_n)) = \frac{1}{\log 2} \log(1 + r_n) = \frac{r_n}{\log 2} + O(r_n^2).$$

Then consider each case (1) and (2) in the corollary using $r_n = c \log\log n / n$ for the (one-sided) ball $[0, r_n]$, and taking in turn $c > \log 2$, followed by $c < \log 2$. When $c = \log 2$, further (additive) error term refinements can be obtained.

3. Towards a dichotomy result for the almost sure growth of maxima

Our aim is to recover versions of the Robbins–Siegmund series critera (1) and (2), as applied to the maximum process $M_n = \max_{k\leq n-1} \phi(f^k)$, where we consider a measure preserving system $(f, \mathcal{X}, \mu)$, and μ is an ergodic measure. The systems we consider include those that can be modelled by a Young tower [50], but in the statement of our results we just require control on the rate of decay of correlations. We make these statements precise as follows.

Definition 3.1. We say that $(f, \mathcal{X}, \mu)$ has decay of correlations in (Banach spaces) $\mathcal{B}_1$ versus $\mathcal{B}_2$, with rate function $\Theta(j) \to 0$ if for all $\varphi_1 \in \mathcal{B}_1$ and $\varphi_2 \in \mathcal{B}_2$ we have

$$\mathcal{C}_j(\varphi_1, \varphi_2, \mu) := \left| \int \varphi_1 \cdot \varphi_2 \circ f^j \, \mathrm{d}\mu - \int \varphi_1 \, \mathrm{d}\mu \int \varphi_2 \, \mathrm{d}\mu \right| \leq \Theta(j) \|\varphi_1\|_{\mathcal{B}_1} \|\varphi_2\|_{\mathcal{B}_2},$$

where $\| \cdot \|_{\mathcal{B}_i}$ denote the corresponding norms on the Banach spaces.

In particular, we consider the L^1 and BV norms of functions $\varphi\colon \mathcal{X} \subset \mathbb{R} \to \mathbb{R}$, defined by

$$\|\varphi\|_1 = \int |\varphi| \, \mathrm{d}\mu,$$
$$\|\varphi\|_{\mathrm{BV}} = \mathrm{var}(\varphi) + \sup(|\varphi|),$$

where $\mathrm{var}(\varphi)$ denotes the total variation of φ. Functions $\varphi\colon \mathcal{X} \subset \mathbb{R} \to \mathbb{R}$ with $\|\varphi\|_{\mathrm{BV}} < \infty$ are called functions of bounded variation.

The first main assumption is the following.

(A1) **Exponential decay of correlations** with respect to notations of Definition 3.1. We assume that $(f, \mathcal{X}, \mu)$ has exponential decay of correlations in Banach spaces $\mathcal{B}_1 = \mathrm{BV}$ versus $\mathcal{B}_2 = L^\infty$.

Next, we define for a sequence $r_n \to 0$ and integer $p \in \mathbb{N}$ the following quantity

$$\Xi_{p,n} \equiv \Xi_{p,n}(r_n) := \sum_{j=1}^{p} \mu\big(B(\tilde{x}, r_n) \cap f^{-j} B(\tilde{x}, r_n)\big). \tag{8}$$

Our second important assumption is the following.

(A2) **Short Return Times estimate.** Let $s, \gamma \in (0,1)$, and suppose $\gamma + s < 1$. Furthermore, suppose that $B(\tilde{x}, r_n)$ is a sequence of balls with $\sum_n \mu(B(\tilde{x}, r_n)) = \infty$ and $\mu(B(\tilde{x}, r_n)) = O(n^{-\sigma})$ for some $\sigma \in (0,1)$ satisfying

$$\sigma > \max\{\frac{1}{2}(1 + \gamma + s), 1 - \gamma, \frac{1}{3}(2 + s), 1 - \frac{s}{2}\}. \tag{9}$$

Then along the sequence $p_n = n^s$, we have

$$\Xi_{p_n,n}(r_n) = O\big(\frac{1}{n^{1+\gamma}}\big). \tag{10}$$

Condition (A1) is known to hold for a wide class of dynamical systems, such as uniformly expanding maps [42, 44], the Gauß map [47, 48], and also certain non-uniformly expanding quadratic maps [49]. In other applications, modified versions of (A1) include taking $\mathcal{B}_1$ as the space of Hölder continuous functions (maintaining $\mathcal{B}_2 = L^\infty$). We will do this on a case-by-case basis.

Condition (A2) gives a restriction on the recurrence properties of $\tilde{x} \in \mathcal{X}$, and for a broad class of systems, this condition can be proved to hold for μ-a.e. $\tilde{x} \in \mathcal{X}$ along the lines of [12, 28, 30]. For readers familiar with extreme value theory, equation (10) is similar to the $D'(u_n)$ condition considered in [14, 43, 45], where u_n plays the role of $\psi(r_n)$. Some of the restrictions on the constants within equation (9) arise through requiring self-consistency of equation (10). Indeed, by exponential decay of correlations (A1) we have $\mu(B(\tilde{x}, r_n) \cap f^{-k} B(\tilde{x}, r_n)) \approx \mu(B(\tilde{x}, r_n))^2$ for $k \gg \log n$. Hence, the bound in equation (10) forces $2\sigma - s > 1 + \gamma$. In Section 7 we discuss the role for the other lower bounds on σ within (9). Thus, condition (A2) mainly applies to control the measure $\mu(B(\tilde{x}, r_n) \cap f^{-k} B(\tilde{x}, r_n))$ for $k = O(\log n)$, i.e. for short return times. In turn this condition is a restriction on the recurrence statistics of $\tilde{x}$ and nearby points. In Section 13 we provide general techniques to verify (A2). For a wide class of dynamical systems we show that (A2) holds for μ-a.e. $x \in \mathcal{X}$. The techniques we discuss build upon and complement arguments used in Collet [12]. Examples include piecewise expanding maps with absolutely continuous invariant measures or more general Gibbs measures, and quadratic maps with Benedicks–Carleson parameters. For the latter, (A1) is also satisfied by Young [49].

As stated, Condition (A2) is mainly applicable for systems satisfying (A1). For systems with polynomial decay of correlations, see [28, 30] for conditions similar to (A2). In this article, we consider mainly systems with exponential decay of correlation. The exception is the Manneville–Pomeau map which we treat in Section 4.1.2. We remark further that the exceptional set of points $\tilde{x}$ for which (A2) fails includes periodic points. We discuss this further in Section 4.

To state the next theorem we consider an interval map f and we recall that a measure ν is called conformal for the non-negative function $g \colon \mathcal{X} \to \mathbb{R}^+$ if for every measurable set $E \in \mathcal{B}$, on which f acts as a measurable isomorphism, we have

$$\nu(f(E)) = \int_E g \, \mathrm{d}\nu.$$

One often refers to $\log g$ as a potential. Furthermore, the transfer operator $\mathscr{L} \colon L^1(\mathcal{X}, \nu) \to L^1(\mathcal{X}, \nu)$ (sometimes called Ruelle operator or Perron–Frobenius operator) is defined by

$$\mathscr{L}\psi(x) = \sum_{f(y)=x} g(y)\psi(y).$$

We often restrict $\mathscr{L}$ to the functions of bounded variation.

Suppose $\lambda > 0$ is the maximal eigenvalue of $\mathscr{L}$ as acting on functions of bounded variations and that h is a non-negative function such that $\mathscr{L}(h) = \lambda h$. Then, if ν is a conformal measure with respect to $1/(\lambda g)$, then

$$\int \psi \, \mathrm{d}\nu = \lambda^{-1} \int \mathscr{L}(\psi) \, \mathrm{d}\nu.$$

We can define a measure μ by

$$\int \psi \, \mathrm{d}\mu = \int \psi h \, \mathrm{d}\nu, \qquad \text{for all continuous } \psi.$$

The measure μ is f-invariant since

$$\int \psi \circ f \,\mathrm{d}\mu = \int \psi \circ f h \,\mathrm{d}\nu = \lambda^{-1} \int \mathscr{L}(\psi \circ f h) \,\mathrm{d}\nu$$
$$= \lambda^{-1} \int \psi \mathscr{L}(h) \,\mathrm{d}\nu = \int \psi h \,\mathrm{d}\nu = \int \psi \,\mathrm{d}\mu,$$

and μ is called a *Gibbs measure* with respect to g. (Replacing g by g/λ we may assume that $\lambda = 1$.)

Theorem 3.2. *Suppose that $f\colon \mathcal{X} \to \mathcal{X}$ is an interval map with an ergodic probability measure μ. Assume that condition (A1) holds. Furthermore, assume that μ is a Gibbs measure which has a density h with respect to a conformal measure, that h is an eigenfunction of a transfer operator with a spectral gap when acting on function of bounded variation, and that h is the unique (up to scalar multiples) eigenfunction of maximal modulus of the eigenvalue. Consider the observable $\phi(x) = \psi(\operatorname{dist}(x, \tilde{x}))$, where $\psi\colon [0,\infty) \to \overline{\mathbb{R}}$ is a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$, and suppose that condition (A2) holds for the sequence of balls $\{B(\tilde{x}, r_n)\}$, $r_n \to 0$ centered at $\tilde{x}$. Moreover, suppose the sequence (r_n) is non-increasing and such that $n \mapsto n\mu(B(\tilde{x}, r_n))$ is non-decreasing. Then we have the following dichotomy:*

(1) If the sequence (r_n) satisfies

$$\sum_{n=1}^{\infty} \mu(B(\tilde{x}, r_n)) e^{-n\mu(B(\tilde{x}, r_n))} < \infty,$$

then

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(\psi(r_n) \le M_n \text{ ev.}\big) = 1.$$

(2) If the sequence (r_n) satisfies

$$\sum_{n=1}^{\infty} \mu(B(\tilde{x}, r_n)) e^{-n\mu(B(\tilde{x}, r_n))} = \infty,$$

then

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(\psi(r_n) \le M_n \text{ ev.}\big) = 0.$$

Remark 3.3. Liverani, Saussol and Vaienti [44] studied a general class of piecewise expanding interval maps and a Gibbs measure μ with respect to a potential. They proved that under some mild regularity conditions and a "covering" condition, that the transfer operator related to the potential has a spectral gap and a unique (up to scaling) eigenfunction associated to the eigenvalue of maximal modulus. A special case is that μ is a measure which is absolutely continuous with respect to Lebesgue measure. We discuss various dynamical system case studies in Section 13, including piecewise differentiable maps satisfying (A1). For these systems we show also that condition (A2) holds for a.e. $\tilde{x}$.

There are systems which do not satisfy the assumptions of Theorem 3.2, but for which we are able to prove a similar result. To state this result, we introduce the following complexity growth condition.

(A3) There exists $K_f \in \mathbb{N}$, such that for all $r \ge 0$ and all x the set $f^{-1}(B(x, r))$ has at most K_f connected components.

Condition (A3) is satisfied by maps with finitely many monotone branches (such as unimodal maps). For these systems the density of μ is not necessarily bounded and hence not in BV as required by Theorem 3.2. We have the following result.

Theorem 3.4. *Suppose that $f\colon \mathcal{X} \to \mathcal{X}$ is an interval map with ergodic probability measure μ, and assume that condition (A1) holds.*

Consider the observable $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$, where $\psi\colon [0, \infty) \to \overline{\mathbb{R}}$ is a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$, and suppose that condition (A2) holds for sequences of balls centered at $\tilde{x}$. Moreover, suppose the sequence (r_n) is non-increasing and such that $n \mapsto n\mu(B(\tilde{x}, r_n))$ is non-decreasing.

We have the following cases.

(1) *Suppose the sequence (r_n) satisfies*

$$\sum_{n=1}^{\infty} \mu(B(\tilde{x}, r_n)) e^{-n\mu(B(\tilde{x}, r_n))} < \infty.$$

Then for any $a > 1$, we have

$$\mu\big(M_n \geq \psi(r_{[\frac{n}{a}]}) \text{ ev.}\big) = 1.$$

In particular, if $\mu\{\phi \geq u_n\} \geq c\frac{\log\log n}{n}$ for some constant $c > 1$, then

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(u_n \leq M_n \text{ ev.}\big) = 1.$$

(2) *Suppose condition (A3) holds, and the sequence (r_n) satisfies*

$$\sum_{n=1}^{\infty} \mu(B(\tilde{x}, r_n)) e^{-n\gamma\mu(B(\tilde{x}, r_n))} = \infty,$$

for some $\gamma > 1$. Then we have

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(\psi(r_n) \leq M_n \text{ ev.}\big) = 0.$$

In particular, if $\mu(\phi \geq u_n) \leq c\frac{\log\log n}{n}$ for some constant $c < 1$, then

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(u_n \leq M_n \text{ ev.}\big) = 0.$$

Remark 3.5. An example fitting this theorem is a quadratic map with Benedicks–Carleson parameter. For such systems it can be proved along the lines of the estimates by Collet [12] that (A2) holds for almost all $\tilde{x}$. See Section 13 for precise statements.

Remark 3.6. The proof of Theorem 3.4 uses a Cauchy-condensation method. The method turns out to be quite versatile in obtaining the lower bound sequence $u_n = \psi(r_n)$ for M_n, but is less applicable for establishing dichotomy results, i.e. to understand when $\nu(M_n \leq u_n \text{ i.o.}) = 1$. Thus to prove case (2) we instead follow a method similar to that used for case (2) of Theorem 3.2. The contrasting bounds obtained from these two theorems are very fine. Indeed, from Theorem 3.2, a lower bound sequence u_n satisfies

$$\mu(X_1 > u_n) \geq \frac{\log\log n}{n} + \frac{c\log\log\log n}{n}, \qquad \text{for some } c > 2.$$

This sequence is a narrow improvement on the range of lower bound sequences implied by Theorem 3.4. Similarly for the sequences that determine when $\mu(\mathcal{H}_{\mathrm{ea}}) = 0$.

In Section 10.1 we deduce the "in particular part of (1)" from Theorem 3.4 (with the usual choice $u_n = \psi(r_n)$), while the "in particular part of (2)" follows by direct calculation.

It is worth to compare these bounds with Corollaries 1.3 and 1.4 in [40]. Imposing a long-term independence property on the shrinking target system they obtain tight conditions on the shrinking rate of the targets so that $\mathcal{H}_{\mathrm{ea}}$ has zero or full measure. In particular, their assumptions are satisfied for specific choices of targets in product systems and Bernoulli shifts. In the case of product systems, [40, Corollary 1.3] yields that the shrinking rate $\mu(B_n) \geq \frac{c \log \log n}{n}$ for some $c > 1$ implies $\mu(\mathcal{H}_{\mathrm{ea}}(\mathbf{B})) = 1$, while $\mu(B_n) \leq \frac{\log \log n}{n}$ for all but finitely many n implies $\mu(\mathcal{H}_{\mathrm{ea}}(\mathbf{B})) = 0$.

In Section 4, we discuss results and examples in the case where the short return time condition (A2) fails. For these examples the maximum process has a non-trivial extremal index $\theta \in (0, 1)$. We also show that the Robbins–Siegmund series criteria fails, and propose more general criteria on what the bounding sequences u_n and v_n should satisfy.

4. Almost sure growth of maxima for processes having an extremal index

In this section we consider again measure preserving dynamical system $(f, \mathcal{X}, \mu)$ on the unit interval, and the maximum process $M_n = \max_{k \leq n-1} \phi \circ f^k$. However, we consider the situation where equations (4) and (5) apply for a non-trivial extremal index $\theta \in (0, 1)$. We show that the Robbins–Siegmund series criteria as stated in (1) and (2) are no longer valid for producing the (almost sure) bounding sequences for the process M_n. We obtain modified series criteria based on inclusion of the parameter θ. Towards the end of this section, we propose a general question on the validity of such series criteria for bounding M_n in the case of general dependent processes, i.e. beyond dynamical systems.

More explicitly, we consider situations where the short return time condition (A2) fails. For dynamical systems, this can happen in the case for a sequence of shrinking targets limiting onto a periodic point. To state our main results, we shall focus on this case. Indeed, for observables of the form $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$ a non-trivial extremal index tends to only arise in these cases, especially for the dynamical systems we consider. However, there are many other mechanisms that can give rise to a non-trivial θ, for an overview see [45].

For an observable maximized at a periodic point $\tilde{x} \in \mathcal{X}$, assumption (A2) can be shown to fail as follows. Suppose $f^p(\tilde{x}) = \tilde{x}$, for some $p \geq 1$. Then

$$\mu(B(\tilde{x}, r) \cap f^{-p}(B(\tilde{x}, r))) > C_p \mu(B(\tilde{x}, r)),$$

where $C_p > 0$ depends on the derivative of f^p at $\tilde{x}$ and the measure μ. (The constant is non-zero, if $(f^p)'(\tilde{x}) < \infty$ and μ is equivalent to Lebesgue measure, at least locally at $\tilde{x}$). From the view of extreme value theory, the maximum process has a distribution governed by a non-trivial extremal index, as described by equations (4) and (5).

The blocking arguments that we use to prove Theorems 3.2–3.4 must be adapted to account for the failure of condition (A2). The relevant modifications are discussed in (for example) [18], and we review the relevant constructions. We introduce the events

$$U(u) = \{X_1 > u\},$$
$$A^{(q)}(u) = U(u) \cap \bigcap_{k=1}^{q} f^{-k}(U(u)^{\complement}).$$

In the case $u := u_n$ we write $U_n = U(u_n)$, and $A_n^{(q)} = A^{(q)}(u_n)$. Define a sequence θ_n via

$$\theta_n = \frac{\mu(A_n^{(q)})}{\mu(U_n)}.$$

In the setting of a dynamical system with a q-periodic point $\tilde{x}$ and an observable $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$, the event $A^{(q)}(u_n)$ gives the points in $U_n = B(\tilde{x}, r_n)$, where $u_n = \psi(r_n)$, that do not return to $B(\tilde{x}, r_n)$ within q iterates. Accordingly, θ_n is the proportion of points in $B(\tilde{x}, r_n)$ that do not return within q iterates.

When $\theta = \lim_{n\to\infty} \theta_n$ exists, then this constant $\theta \in (0,1]$ takes the role as the extremal index. We make the following convergence assumption.

(A4) There exists $\hat{\sigma} > 0$ such that

$$|\theta_n - \theta| = O(n^{-\hat{\sigma}}),$$

and $\theta \neq 0$.

Assumption (A4) is an assumption on the local property of the dynamical system at $\tilde{x}$. Verification depends on assumptions of the regularity of the invariant density and derivative of f at $\tilde{x}$. It is easy to see that Assumption (A4) is valid when $\tilde{x}$ is a hyperbolic repelling periodic point for a piecewise (linear) expanding map, and μ is Lebesgue measure. In these cases, for the limit we get

$$\theta = 1 - \frac{1}{|(f^q)'(\tilde{x})|}. \tag{11}$$

Indeed, if $f(x)$ is the doubling map $f(x) = 2x \mod 1$ on $[0,1]$, then we get the exact formula

$$\theta_n = \theta = 1 - \frac{1}{2^q},$$

for all n sufficiently large. The result of equation (11) also holds in greater generality, see [17]. We remark that $\theta \neq 0$ is assumed in (A4). In Section 4.1.3 we consider an example where $\theta = 0$.

For a broad class of non-uniformly expanding dynamical systems, see [17] for examples where θ is computed together with establishing the corresponding limit law for the maxima. We have the following result.

Theorem 4.1. *Suppose that $f\colon \mathcal{X} \to \mathcal{X}$ is an interval map with an ergodic probability measure μ. Assume that condition (A1) holds. Furthermore, assume that μ is a Gibbs measure which has a density h with respect to a* *conformal measure, that h is an eigenfunction of a transfer operator with a spectral gap when acting on function of bounded variation, and that h is the unique (up to scalar multiples) eigenfunction of maximal modulus of the eigenvalue.*

Consider the observable $\phi(x) = \psi(\mathrm{dist}(x,\tilde{x}))$, where $\psi\colon [0,\infty) \to \overline{\mathbb{R}}$ is a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$, and $\tilde{x}$ is a hyperbolic periodic point of period q, and $|(f^q)'(\tilde{x})| < \infty$. Consider a sequence of balls of radii $r_n \to 0$, each centered at $\tilde{x}$. Suppose that (A4) holds at $\tilde{x}$, and that the sequence (r_n) is non-increasing and such that $n \mapsto n\mu(B(\tilde{x},r_n))$ is non-decreasing. Then we have the following dichotomy:

(1) *If the sequence (r_n) satisfies*

$$\sum_{n=1}^{\infty} \mu(B(\tilde{x},r_n))e^{-n\theta\mu(B(\tilde{x},r_n))} < \infty,$$

then

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(\psi(r_n) \leq M_n \text{ ev.}\big) = 1.$$

(2) *If the sequence (r_n) satisfies*

$$\sum_{n=1}^{\infty} \mu(B(\tilde{x},r_n))e^{-n\theta\mu(B(\tilde{x},r_n))} = \infty,$$

then

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(\psi(r_n) \leq M_n \text{ ev.}\big) = 0.$$

In both cases θ denotes the corresponding extremal index.

We prove Theorem 4.1 in Section 11. See also [17, Section 4] for similar discussions in the distributional convergence cases. The following result concerns the eventual lower bound for the maximum process, and can be contrasted to Theorem 3.4.

Theorem 4.2. *Suppose that $f\colon \mathcal{X} \to \mathcal{X}$ is an interval map with ergodic probability measure μ, and assume that condition (A1) holds. Consider the observable $\phi(x) = \psi(\mathrm{dist}(x,\tilde{x}))$, where $\psi\colon [0,\infty) \to \overline{\mathbb{R}}$ is a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$, and $\tilde{x}$ is a hyperbolic periodic point of period q, and $|(f^q)'(\tilde{x})| < \infty$. Consider a sequence of balls of radii $r_n \to 0$, each centered at $\tilde{x}$. Suppose that (A4) holds at $\tilde{x}$, and that the sequence (r_n) is non-increasing and such that $n \mapsto n\mu(B(\tilde{x},r_n))$ is non-decreasing. We have the following cases.*

(1) *Suppose the sequence (r_n) satisfies*

$$\sum_{n=1}^{\infty} \mu(B(\tilde{x},r_n))e^{-n\theta\mu(B(\tilde{x},r_n))} < \infty.$$

Then for any $a > 1$, we have

$$\mu\big(M_n \geq \psi(r_{[\frac{n}{a}]}) \text{ ev.}\big) = 1.$$

In particular, if $\mu(\phi \geq u_n) \geq c\theta^{-1}\frac{\log\log n}{n}$ for some constant $c > 1$, then

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(u_n \leq M_n \text{ ev.}\big) = 1.$$

(2) *Suppose condition (A3) holds, and the sequence* (r_n) *satisfies*

$$\sum_{n=1}^{\infty} \mu(B(\tilde{x}, r_n)) e^{-n\theta\gamma\mu(B(\tilde{x},r_n))} = \infty,$$

for some $\gamma > 1$. *Then we have*

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(\psi(r_n) \le M_n \ ev.\big) = 0.$$

In particular, if $\mu(\phi \ge u_n) \le c\theta^{-1}\frac{\log\log n}{n}$ *for some constant* $c < 1$, *then*

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(u_n \le M_n \ ev.\big) = 0.$$

The proof of Theorem 4.2 is also given in Section 11. The method of proof uses Cauchy-condensation techniques.

4.1. **Examples and general discussion of dichotomy criteria.** In this section we consider further examples, including a dichotomy criterion applied to a non-dynamical example. We also consider an example where the extremal index is zero.

4.1.1. *An example of a stochastic process satisfying the dichotomy for a non-trivial* $\theta \in (0, 1)$. We consider a stationary stochastic process which has a non-trivial extremal index. The mechanism giving rise to the extremal index here is different to the periodic phenomena arising in the dynamical systems.

Suppose (X_n) is an i.i.d. process with distribution function $F_X(x) = P(X < x) = 1 - x^{-1}$. We let $\overline{F}_X(x) = 1 - F_X(x)$ (the tail distribution). Consider the process $Y_n = \max(X_n, aX_{n+1})$, where $a \ge 1$ is fixed. We claim the following:

(1) The extremal index for the process $M_n^Y = \max_{k \le n} Y_k$ is given by $\theta = a/(a + 1)$.
(2) Dichotomy criteria, as consistent with the conclusion of Theorem 4.1 hold.

First of all (since $a \ge 1$) we have

$$M_n^Y = \max_{k \le n} Y_k = \max\{X_1, aX_2, aX_3, \dots, aX_{n+1}\}.$$

Now we compute the extremal index for the process Y_n. Let $\overline{F}_Y(y) := 1 - P(Y_1 < y)$. Then

$$\begin{aligned}\overline{F}_Y(y) &= 1 - P(Y_1 < y)\\ &= 1 - P(X_1 < y, X_2 < y/a)\\ &= 1 - (1 - y^{-1})(1 - ay^{-1})\\ &= \frac{a+1}{y} + O(y^{-2}).\end{aligned}$$

Given $\tau > 0$ let $u_n = (a+1)n/\tau$. Then this sequence u_n satisfies $nP(Y_1 > u_n) \to \tau$. Now consider the distribution for the process M_n^Y. We have

$$\begin{aligned} P(M_n^Y < u_n) &= P(X_1 < u_n)P(X_2 < u_n/a)\dots P(X_{n+1} < u_n/a) \\ &= \left(1 - \frac{1}{u_n}\right)\left(1 - \frac{a}{u_n}\right)^n \\ &= \left(1 - \frac{\tau}{(a+1)n}\right)\left(1 - \frac{a\tau}{(a+1)n}\right)^n \\ &\to e^{-\theta\tau}, \qquad n \to \infty, \end{aligned}$$

with $\theta = a/(a+1)$. Consider almost sure bounds for the maxima of (X_n) via general sequences $\tilde{u}_n$ and $\tilde{v}_n$, so that $\tilde{v}_n < M_n^X < \tilde{u}_n$ (almost surely). Consider also the intermediate sequence $\tilde{w}_n$ with $P(M_n^X > \tilde{w}_n \text{ i.o.}) = P(M_n^X < \tilde{w}_n \text{ i.o.}) = 1$. We assume that these sequences satisfy monotonicity constraints as consistent with assumptions of Theorem 4.1. By definition of the process (Y_n) we obtain almost surely, for all n sufficiently large

$$M_n^Y = \max\{X_1, aX_2, \dots aX_{n+1}\} \leq a\tilde{u}_{n+1}, \tag{12}$$

$$M_n^Y \geq \max\{aX_2, \dots aX_{n+1}\} \geq a\tilde{v}_n, \tag{13}$$

and by similar analysis, almost surely there are infinite subsequences $\tilde{w}_{n_k}$ with

$$M_{n_k}^Y > a\tilde{w}_{n_k}, \qquad M_{n_k}^Y < a\tilde{w}_{1+n_k}. \tag{14}$$

Using the explicit regularity of the tails $\overline{F}_X(x)$, and $\overline{F}_Y(y)$, we see that divergence or convergence of the sum

$$\sum_n \overline{F}_X(z_n)e^{-n\overline{F}_X(z_n)} \tag{15}$$

is equivalent to divergence or (resp.) convergence of the sum

$$\sum_n \overline{F}_Y(az_n)e^{-\frac{an}{a+1}\overline{F}_Y(az_n)}.$$

We now apply this to $z_n = \tilde{v}_n$ and $z_n = \tilde{w}_n$. A simple index relabelling of the series in equation (15) shows that convergence/divergence is unaffected by the index translation $n \to n \pm 1$. Indeed, since $e^{\overline{F}_X(z_n)} = e^{o(1)}$, divergence/convergence of equation (15) is equivalent to divergence/convergence of

$$\sum_n \overline{F}_X(z_{n-1})e^{-n\overline{F}_X(z_{n-1})}.$$

This allows us to replace the relevant upper bound sequence $\tilde{w}_{n_k+1}$ in equation (14) by $\tilde{w}_{n_k}$. We remark that the upper bound sequence $\tilde{u}_{n+1}$ as appearing in equation (12) can also be replaced by $\tilde{u}_n$. This follows using the First Borel Cantelli Lemma, monotonicity of u_n, and a simple index translation for the (convergent) sum $\sum_n \overline{F}_X(\tilde{u}_n)$. This concludes the example.

4.1.2. *The Manneville–Pomeau map.* In this section we consider how the method of inducing allows us to extend the conclusions of Theorem 3.2 and Theorem 4.1 to a wider range of examples, e.g. to dynamical systems whose transfer operator does not have a spectral gap. We illustrate using the Manneville–Pomeau map $f\colon [0,1] \to [0,1]$ given by

$$f(x) = \begin{cases} x(1+2^a x^a) & \text{if } 0 \le x < 1/2, \\ 2x-1 & \text{if } 1/2 \le x \le 1, \end{cases} \tag{16}$$

with $a \in (0,1)$. In the following we let $A = [1/2, 1]$, but the construction we describe extends to any interval of the form $[\epsilon, 1]$ for $\epsilon > 0$. See [21, 36]. Consider the first return map $\hat{f}\colon A \to A$ given by $\hat{f}(x) = f^{R(x)}(x)$, with $R(x) = \inf\{\, n \ge 1 : f^n(x) \in A \,\}$, (and $x \in A$). The map $\hat{f}$ preserves an absolutely continuous invariant measure $\hat{\mu}$, with density in BV, and has exponential decay of correlations of BV against L^1 [36, section 4.2]. Thus Theorems 3.2 and 4.1 apply to $(\hat{f}, \hat{\mu})$, for observables of the form $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$ with $\tilde{x} \in A$. Let $\widehat{M}_n(x) = \max_{1 \le k \le n} \phi(\hat{f}^{k-1}(x))$, and (as before) $M_n(x) = \max_{1 \le k \le n} \phi(f^{k-1}(x))$. Then we have

$$M_n(x) = \max_{1 \le j \le k(n,x)} \phi(\hat{f}^{j-1}(x)) = \widehat{M}_{k(n,x)}(x),$$

where $x \in A$, and $k(n,x)$ satisfies

$$\sum_{j=1}^{k(n,x)-1} R(\hat{f}^{j-1}(x)) \le n < \sum_{j=1}^{k(n,x)} R(\hat{f}^{j-1}(x)).$$

Since f preserves an absolutely continuous invariant measure μ, we deduce that for μ-almost all $x \in A$ that $k(n,x)/n \to \mu(A)$. We get the following result.

Proposition 4.3. *Suppose $f\colon [0,1] \to [0,1]$ is given by equation* (16)*. Let $A = [1/2, 1]$ and $a \in (0,1)$. Let $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$, with $\psi\colon [0,\infty) \to \overline{\mathbb{R}}$ a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$. Then we have the following result for almost every $\tilde{x} \in A$.*

1. *Suppose that the sequence (r_n) is non-increasing and satisfies $\mu(B(\tilde{x}, r_n)) > c \log\log n/n$ for some $c > 1$, then*
$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(\psi(r_n) \le M_n \text{ ev.}\big) = 1.$$
2. *Suppose that the sequence (r_n) is non-increasing and satisfies $\mu(B(\tilde{x}, r_n)) < c \log\log n/n$ for some $c < 1$, then*
$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu\big(\psi(r_n) \le M_n \text{ ev.}\big) = 0.$$

Remark 4.4. We have stated the result only for $a \in (0,1)$. It is possible to consider also $a \ge 1$ where the f-invariant ergodic measure μ is no longer finite. However, we do not get significant improvements over results obtained in [36] due to the (almost sure) fluctuations in $R(x)$.

Proof. By Proposition 13.4, condition (A2) holds for $\hat{f}$ and μ-a.e. $\tilde{x} \in [0,1]$. Thus $\hat{f}$ satisfies the assumptions of Theorems 3.2.

The proof then follows step by step that of [36, Theorem 2]. The only modification is that bounding sequences for $M_n(x)$ are determined via Theorem 3.2 as applied to the map $\hat{f}$ to bound $\widehat{M_k}(x)$. Here $k \equiv k(n,x)$ with $k(n,x)/n \to \mu(A)$ almost surely. □

4.1.3. *An example with extremal index $\theta = 0$.* We consider again the Manneville–Pomeau map given by equation (16), and take an observable function of the form $\phi(x) = \psi(d(x,0))$, thus maximized at the point $\tilde{x} = 0$. It is shown in [19] that the distribution for the maxima follows a degenerate extreme value law with extremal index equal to zero (under a scaling sequence given by (4)). We obtain the following almost sure result.

Proposition 4.5. *Suppose $f\colon [0,1] \to [0,1]$ is given by equation (16) with $a \in (0,1)$, and $\phi(x) = \psi(\mathrm{dist}(x,0))$, with $\psi\colon [0,\infty) \to \overline{\mathbb{R}}$ a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$. Then we have the almost sure result $\mu(v_n \le M_n \le u_n \text{ ev.}) = 1$, where u_n and v_n satisfy*

$$\mu(X_1 > v_n) \ge \left(\frac{c \log\log n}{n}\right)^{1-a}, \quad \mu(X_1 > u_n) \le \left(\frac{1}{n(\log n)^2}\right)^{1-a},$$

with $c > c_0$, for some $c_0 > 0$ (depending on the density of μ).

We prove Proposition 4.5 in Section 11.2. This result also refines estimates obtained in [21, 29], especially on the lower bound sequence v_n. The upper bound sequence (as stated) can be refined easily using the First Borel–Cantelli Lemma. This result illustrates that when $\theta = 0$, we can expect non-standard growth rates for the maxima. The exponent $1-a$ arises due to the presence of the non-hyperbolic fixed point $\tilde{x} = 0$.

4.1.4. *On a general dichotomy criteria.* Within this section we have exhibited processes giving rise to a non-trivial extremal index. These processes are created by using underlying periodic phenomena of the dynamical system process. A non-trivial index can arise also through other non-periodic mechanisms within dynamical systems, see [2, 7]. More broadly, clustering can arise in more general process settings, see [14, 45]. It is therefore natural to ask whether the conclusion of Theorem 4.1 is applicable in wider scenarios. We have given in Section 4.1.1 a (non-dynamical) example to illustrate that this is still the case. However, a general criteria is yet to be fully developed on determining the sequences u_n for which a zero–one law applies to $\mu(M_n \le u_n \text{ i.o.})$. We consider the following question.

Problem 4.6. For what class of stationary stochastic processes (X_n) does the following hold? There exists a constant $\theta \in (0,1]$ such that

(1) If u_n is such that
$$\sum_{n=1}^{\infty} \mu(X_1 > u_n) = \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \mu(X_1 > u_n) e^{-n\theta\mu(X_1 > u_n)} < \infty$$
then $\mu(M_n \ge u_n \text{ ev.}) = 1$;

(2) If u_n is such that
$$\sum_{n=1}^{\infty} \mu(X_1 > u_n) = \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \mu(X_1 > u_n) e^{-n\theta\mu(X_1 > u_n)} = \infty$$

then $\mu(M_n \le u_n \text{ i.o.}) = \mu(M_n \ge u_n \text{ i.o.}) = 1$.

For the i.i.d. case, these items are both valid for $\theta = 1$. However a full classification of processes (X_n) which satisfy these criteria for given $\theta \in (0,1)$ is unknown. For certain maximum processes with extremal index $\theta \in (0,1)$ we have shown that these dichotomy conditions apply.

5. Almost sure bounds for M_n for non-uniformly hyperbolic systems

In this section we consider almost sure bounds on M_n for a wide class of hyperbolic systems. These include systems where $(f, \mathcal{X}, \mu)$ is modelled by a Young tower [50]. We emphasize that the results in this section are valid for some higher dimensional systems, while results of previous sections are for interval maps. For higher dimensional systems, e.g. such as those that admit Sinai–Ruelle–Bowen measures, obtaining distributional convergence of the maxima requires the blocking arguments used for one-dimensional systems to be significantly modified. This includes having additional regularity constraints placed on μ (depending upon the strength of result obtained). With the present techniques available, we obtain results towards Case (1) of Theorems 3.4 and 4.2. We consider the following assumption on the distributional convergence.

(A5) Given constants $\sigma_1, \sigma_2 > 1$, then for every $r \in (0,1)$ there is a set $\mathcal{M}_r$, with $\mu(\mathcal{M}_r) \le C_1 |\log r|^{-\sigma_1}$, such that for all $\tilde{x} \notin \mathcal{M}_r$,

$$\left| \mu\left(x : \sum_{k=0}^{t/\mu(B_r(\tilde{x}))} \mathbb{1}_{B_r(\tilde{x})}(f^k(x)) = 0 \right) - e^{-\theta t} \right| \le C_2 |\log r|^{-\sigma_2}, \tag{17}$$

for all $t \ge 0$.

Assumption (A5) is recognised as an approximate *exponential law* for entrance times to shrinking balls. For certain non-uniformly hyperbolic systems modelled by Young towers, (A5) is shown to hold, see for example [9, 31] (in the case $\theta = 1$), where a more general Poisson laws result can also be obtained. Examples of such systems include those with Axiom A attractors, and the Hénon map family for Benedicks–Carleson parameters [5]. To bring assumption (A5) in line with distribution results for maxima M_n, consider the observable function $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$, and a sequence $r_n \to 0$. Set $u_n = \psi(r_n)$ and $t \equiv t_n = n\mu(B(\tilde{x}, r_n))$. Here we allow the possibility that $t_n \to 0$ or $t_n \to \infty$. Then equation (17) becomes

$$\left| \mu\big(x : M_n(x) < u_n \big) - e^{-n\theta\mu(B(\tilde{x}, r_n))} \right| \le C_2 |\log r_n|^{-\sigma_2}. \tag{18}$$

This has similarities to the results obtained in Section 7 for one-dimensional systems, in particular Corollary 7.3. However, the approximation of (18) is not uniform over the ball center $\tilde{x}$. In order to mirror Corollary 7.3, we require $\tilde{x} \notin \mathcal{M}_{r_n}$ for all n, and clearly this condition depends on the sequence (r_n). Further arguments are therefore required to avoid the existence of an infinite subsequence (r_{j_k}) for which $\tilde{x} \in \mathcal{M}_{r_{j_k}}$. For hyperbolic systems, the presence of the set $\mathcal{M}_r$ arises from the regularity assumptions (or lack thereof) placed on the measure μ. These issues are discussed in [9, 30, 31]. When stronger regularity properties are known (or assumed) about the measure μ, then it can be shown that $\mu(\limsup_{r\to 0}\mathcal{M}_r) = 0$. This applies for certain uniformly hyperbolic systems and billiard models, see [13, 8, 24, 46].

To state our result, we also need existence of a local dimension $d_\mu(\tilde{x})$ at $\tilde{x}$. This is defined to be the limit

$$d_\mu(\tilde{x}) = \lim_{r\to 0}\frac{\log\mu(B(\tilde{x},r))}{\log r},$$

whenever this limit exists. For a wide range of hyperbolic systems, the value $d_\mu(\tilde{x})$ exists for μ-a.e. $\tilde{x}\in\mathcal{X}$, see [4].

Theorem 5.1. *Suppose $(f,\mathcal{X},\mu)$ is a measure preserving system, and (A5) holds. Given $\tilde{x}\in\mathcal{X}$, let $\phi(x) = \psi(\mathrm{dist}(x,\tilde{x}))$, where $\psi\colon [0,\infty)\to\overline{\mathbb{R}}$ is a decreasing continuous function with $\psi(y)\to\infty$ as $y\to 0$. Suppose that $B(\tilde{x},r_n)$ are nested balls such that $\mu(B(\tilde{x},r_n))\geq c\theta^{-1}\log\log n/n$ for some $c>1$, and that the local dimension $d_\mu(\tilde{x})$ exists μ-almost everywhere. Then for μ-a.e. $\tilde{x}\in\mathcal{X}$ and the sequence $u_n=\psi(r_n)$ we have $\mu(M_n\geq u_n$ ev.$) = 1$, that is, $\mu(\mathcal{H}_{\mathrm{ea}}(\mathbf{B})) = 1$.*

We make several remarks on the proof and scope of this result. For the proof of the result, we by-pass the influence of the set $\mathcal{M}_r$ to obtain a result similar to Corollary 7.3. We can then apply the Cauchy-condensation method used for proving Theorem 3.4. With the current techniques available this is the best we can achieve. The arguments used within Section 9 cannot easily be adapted and new ideas are needed.

Within Theorem 5.1 we require $\sigma_1,\sigma_2>1$. However, in certain applications the possibility $\sigma_1,\sigma_2\in(0,1)$ can arise, see [31]. In this case, we get weaker bounds on sequence $u_n=\psi(r_n)$, namely having the requirement $\mu(X_1 > u_n)\geq e^{(\log n)^{\gamma'}}/n$ for some $\gamma'\in(0,1)$, see Section 12.

A further remark is that having an assumption on the existence of a local dimension can be weakened. From the proof, we generally require quantitative bounds on the decay of $\mu(B(\tilde{x},r_n))$ along certain sequences $r_n\to 0$.

5.1. **Intermediate growth rate sequences for maxima.** In this section we consider non-decreasing sequences (u_n) for which the following statement applies

$$\mu(M_n > u_n \text{ i.o.}) > 0 \qquad\text{and}\qquad \mu(M_n\leq u_n \text{ i.o.}) > 0.$$

Clearly the dichotomy results obtained in e.g. Theorems 3.2 and 4.1 fully describe these sequences. However, in the case of the dynamical systems for which Theorem 5.1 applies, we can obtain partial results using ergodicity of the dynamical system. Since (u_n) is non-decreasing, we have that $\{X_n > u_n \text{ i.o.}\}$ is invariant mod μ. Moreover, if the set $\{\tilde{x}\}$ has zero measure, we have $\{M_n > u_n \text{ i.o.}\} = \{X_n > u_n \text{ i.o.}\}$ mod μ. It then follows by ergodicity that if $\mu(M_n > u_n \text{ i.o.}) > 0$, then $\mu(M_n > u_n \text{ i.o.}) = 1$. On the other hand, $\mu(M_n\leq u_n \text{ i.o.}) > 0$ gives $\mu(\mathcal{H}_{\mathrm{ea}}(\mathbf{B})) = \mu(M_n > u_n \text{ ev.}) < 1$. Thus, $\mu(M_n > u_n \text{ ev.}) = 0$ by ergodicity and invariance mod μ of $\mathcal{H}_{\mathrm{ea}}(\mathbf{B})$ (see [36, Lemma 1] or [34, Lemma 2.4]). This yields $\mu(M_n\leq u_n \text{ i.o.}) = 1$. Hence, both of the measures above will be 1 (if they are positive). We state the following result whose proof is similar to that of [36, Proposition 2].

Theorem 5.2. *Suppose that $(f, \mathcal{X}, \mu)$ is an ergodic measure preserving system satisfying (A5). Consider the observable $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$ with $\psi\colon [0, \infty) \to \overline{\mathbb{R}}$ a decreasing continuous function with $\psi(y) \to \infty$ as $y \to 0$. Suppose that $B(\tilde{x}, r_n)$ are such that $\mu(B(\tilde{x}, r_n)) \leq c/n$, for some $c > 0$. For the sequence $u_n = \psi(r_n)$, we have*

$$\mu(M_n \leq u_n \ i.o.) = 1,$$

that is,

$$\mu(\mathcal{H}_{\mathrm{ea}}(\mathbf{B})) = \mu(M_n > u_n \ ev.) = 0.$$

Proof. We may assume that $\mu(B(\tilde{x}, r_n)) = c/n$, and hence that (u_n) is non-decreasing. Assumption (A5) yields that (18) holds true. Hereby, we get $\mu(M_n \leq u_n) \to e^{-c'} > 0$ for some $0 < c' < \infty$. Furthermore, we recall that

$$\mu(\mathcal{H}_{\mathrm{ea}}) = \mu(M_n > u_n \text{ ev.}) = \mu\left(\bigcup_{i=1}^{\infty} \bigcap_{n=i}^{\infty} \{M_n \leq u_n\}^{\complement} \right).$$

We observe that $\mu(M_n > u_n \text{ ev.}) = \lim_{i \to \infty} \mu(\bigcap_{n=i}^{\infty} \{M_n \leq u_n\}^c) \leq 1 - e^{-c'} < 1$ by nestedness. By ergodicity and invariance mod μ of $\mathcal{H}_{\mathrm{ea}}$ [36, Lemma 1], we conclude the statement. □

In the theorem above, it is possible that we have $\mu(M_n \leq u_n \text{ ev.}) = 1$. However if $\sum_n \mu(B(\tilde{x}, r_n)) = \infty$, and $B(\tilde{x}, r_n)$ is a dynamical Borel–Cantelli sequence, then we instead have $\mu(M_n > u_n \text{ i.o.}) = 1$. For a wide class of dynamical systems, and μ-typical $\tilde{x}$ this property is known to hold, see [1, 10, 25, 26, 29].

6. Overview of the proofs and the blocking argument

Here we give an overview of our proofs. We first use an argument by Galambos: From a dynamical Borel–Cantelli Lemma (for instance [1, 10, 35]) we get from $\sum \mu(B(\tilde{x}, r_n)) = \infty$ that almost surely $X_n \geq u_n$ infinitely often, and hence that $M_n \geq u_n$ holds infinitely often almost surely. Therefore the set $\{M_n < u_n \text{ ev.}\}$ has measure zero. We obtain

$$\begin{aligned} \mu(M_n < u_n \text{ i.o.}) &= \mu(\{M_n < u_n \text{ i.o.}\} \setminus \{M_n < u_n \text{ ev.}\}) \\ &= \mu\big(M_n < u_n \text{ and } M_{n+1} \geq u_{n+1} \text{ i.o.}\big) \\ &= \mu\big(M_n < u_n \text{ and } X_{n+1} \geq u_{n+1} \text{ i.o.}\big), \end{aligned}$$

since if $M_n < u_n$ for infinitely many n, but not for all large enough n, then there must be infinitely many n for which $M_n < u_n$ and $M_{n+1} \geq u_{n+1}$. (Those n may be a strict subset of the k such that $M_k < u_k$.) We will use this equality to prove that $\mu(M_n < u_n \text{ i.o.}) = 0$ and hence that $\mu(M_n \geq u_n \text{ ev.}) = 1$. The idea is to use that for $l < n$

$$\{M_n < u_n \text{ and } X_{n+1} \geq u_{n+1}\} \subset \{M_l < u_n \text{ and } X_{n+1} \geq u_{n+1}\},$$

and if $n - l$ is large, then

$$\begin{aligned} \mu\big(M_n < u_n \text{ and } X_{n+1} \geq u_{n+1}\big) &\leq \mu\big(M_l < u_n \text{ and } X_{n+1} \geq u_{n+1}\big) \\ &\approx \mu(M_l < u_n) \mu(X_{n+1} \geq u_{n+1}), \end{aligned} \tag{19}$$

and this will be made precise using decay of correlation estimates. Then, Proposition 6.1 below is used to estimate $\mu(M_l < u_n)$. This results in an estimate on $\mu(M_n < u_n \text{ and } X_{n+1} \geq u_{n+1})$. Using this estimate, it is shown that $\sum \mu(B(\tilde{x}, r_n)) \exp(-n\mu(B(\tilde{x}, r_n))) < \infty$ implies that

$$\sum_{n=1}^{\infty} \mu(M_n < u_n \text{ and } X_{n+1} \geq u_{n+1}) < \infty$$

and this implies by Borel–Cantelli that

$$\mu(M_n < u_n \text{ and } X_{n+1} \geq u_{n+1} \text{ i.o.}) = 0.$$

Hence $\mu(M_n < u_n \text{ i.o.}) = 0$ and $\mu(M_n \geq u_n \text{ ev.}) = 1$. In this way we obtain the proof of Theorem 3.2.

For some systems, it is difficult to get a good enough error bound in the approximation

$$\mu(M_l < u_n \text{ and } X_{n+1} \geq u_{n+1}) \approx \mu(M_l < u_n)\mu(X_{n+1} \geq u_{n+1})$$

which was used in (19). For such systems we use instead the estimate

$$\mu(M_n < u_n \text{ and } X_{n+1} \geq u_{n+1}) \leq \mu(M_n < u_n).$$

In the end this only leads to a slightly weaker result. This is the path taken to prove Theorem 3.4.

We will now explain how the so-called blocking argument is used to estimate $\mu(M_l < u)$. For general stochastic processes see [14, 43]. Relevant to dynamical systems, the approach we describe is adapted from [12].

We have

$$\{M_l < u\} = \bigcap_{k=1}^{l} \{X_k < u\} \subset \bigcap_{k \in I_l} \{X_k < u\},$$

where $I_l \subset \{1, 2, \ldots, l\}$. We let I_l consist of q blocks of p consecutive numbers in $\{1, 2, \ldots, l\}$, each block separated by t numbers. Writing $I_l = \bigcup_{j=1}^{q} J_j$ where J_j are the blocks, we have

$$\{M_l < u\} \subset \bigcap_{j=1}^{q} \bigcap_{k \in J_j} \{X_k < u\}.$$

The measure of $\bigcap_{k \in J_j} \{X_k < u\}$ is approximated by $1 - p\mu(X_1 \geq u) = 1 - p\mu(\phi \geq u)$ and in this way we can estimate $\mu(M_l < u)$ by $(1 - p\mu(\phi \geq u))^q$. The error obtained by this estimate is expressed in the following proposition by Collet.

Proposition 6.1 (Collet)**.** *Suppose that μ is an ergodic probability measure. Let $s \in (0, \frac{1}{2}]$. With $l = qp + r$, $p = [l^s]$, $0 \leq r < p$ and l large, we have for any u and $t \in \mathbb{N}$ that*

$$\left|\mu(M_l < u) - (1 - p\mu(\phi \geq u))^q\right| \leq \sum_{j=1}^{q} |1 - p\mu(\phi \geq u)|^{q-j} \Gamma_j,$$

where Γ_j is given by

$$\Gamma_j = \left|p\mu(\phi \geq u)\mu(M_{(j-1)(p+t)} < u) - \Sigma_j\right| + t\mu(\phi \geq u) + 2p \sum_{k=1}^{p} \mathsf{E}(\mathbb{1}_{\phi \geq u} \mathbb{1}_{\phi \geq u} \circ f^k)$$

and

$$\Sigma_j = \sum_{k=1}^{p} \mathsf{E}(\mathbb{1}_{\phi\geq u}\mathbb{1}_{M_{(j-1)(p+t)}<u} \circ f^{p+t-k}),$$

where $\mathsf{E}(\cdot)$ *denotes expectation.*

Proposition 6.1 is only interesting if $1 - p\mu(\phi \geq u) > 0$. As we will see, this will be the case in all of our applications of the proposition. The proof of Proposition 6.1 is by Collet [12]. For completeness, we include it in the appendix.

7. Application of Proposition 6.1 and preliminary estimates

In this section we collect several key estimates that we use for proving the main results. We start with an immediate consequence of assumptions (A1) and (A3), where the decay rate is given by $\Theta(j) = \exp(-\tau j)$. This result will be useful for proving Theorem 3.4.

Lemma 7.1. *Let $f\colon \mathcal{X} \to \mathcal{X}$ be an interval map with an ergodic probability measure μ. Assume that conditions (A1) and (A3) hold. Then there are positive constants c_1, K such that for $l \leq n$ we have*

$$\mu(M_n \leq u_n \textit{ and } X_{n+1} > u_{n+1}) \leq \mu(M_l \leq u_n)\mu(X_{n+1} > u_{n+1}) + c_1 K^l e^{-\tau n}.$$

Proof. Let $\varphi_1(x) = \mathbb{1}_{\{M_l<u_n\}}(x)$ and $\varphi_2(x) = \mathbb{1}_{\{X_1>u_{n+1}\}}$. We estimate the BV-norm of $\varphi_1(x)$. Since for any interval A, $f^{-1}(A)$ has at most K_f connected components (by (A3)), it follows that the BV-norm of $\varphi_1(x)$ is bounded by K^l, for some constant K.

Using decay of correlations, we get that

$$\begin{aligned}\mu(M_n \leq u_n \text{ and } X_{n+1} > u_{n+1}) &\leq \int \varphi_1\varphi_2 \circ f^n \,\mathrm{d}\mu \\ &\leq \int \varphi_1 \,\mathrm{d}\mu \int \varphi_2 \,\mathrm{d}\mu + Ce^{-\tau n}\|\varphi_1\|_{\mathrm{BV}}\|\varphi_2\|_\infty \\ &\leq \mu(M_l \leq u_n)\mu(X_{n+1} > u_{n+1}) + c_1 K^l e^{-\tau n}. \qquad \square\end{aligned}$$

In the next step, we use Collet's blocking argument from Proposition 6.1 and the assumptions (A1) and (A2) to obtain an estimate on $\mu(M_l < u_n)$. The following lemma and subsequent corollaries will be used in the proof of most of the Theorems stated within Sections 3 and 4. With a slight change of notation, for integers $p, n \geq 1$ we take

$$\Xi_{p,n} \equiv \Xi_{p,n}(u_n) := \sum_{j=1}^{p} \mu(\phi > u_n, \phi \circ f^j > u_n).$$

This is consistent with the notation of (A2). That is for the sequence r_n as defined in (A2), we have $u_n = \psi(r_n)$. Recall that the observable function is $\phi(x) = \psi(\mathrm{dist}(x, \tilde{x}))$.

Lemma 7.2. *Let $f\colon \mathcal{X} \to \mathcal{X}$ be an interval map with an ergodic probability measure μ. We assume (A1) with rate function $\Theta(j) = \exp(-\tau j)$ and that $\mu(\phi \geq u_n) = O(n^{-\sigma})$ for some $\sigma > \frac{1}{2}$. Then there are constants $C, c_2 > 0$*

and $s \in (0, \frac{1}{2}]$*, such that for sufficiently large* $l \leq n$ *written as* $l = qp + r$ *with* $p = [l^s]$*,* $0 \leq r < p$*, and* $t = [\log(l^{\frac{2}{\tau}})]$ *we have*

$$\mu(M_l < u_n) \leq e^C \exp(-l\mu(\phi \geq u_n)) + qt\mu(\phi \geq u_n) + 2n\Xi_{p,n} + \frac{c_2}{l^{1+s}}. \tag{20}$$

Proof. We start by estimating Γ_j in Proposition 6.1. First of all we note that $p \leq n$. We have $\|\mathbb{1}_{\phi \geq u_n}\|_{\mathrm{BV}} \leq 3$ since $\mathbb{1}_{\phi \geq u_n}$ is the characteristic function of an interval. Hence, by the decay of correlation alone, we have

$$\begin{aligned} &\left|\Sigma_j - p\mu(\phi \geq u_n)\mu(M_{(j-1)(p+t)} < u_n)\right| \\ &= \Bigg|\sum_{k=1}^{p} \Big(\mathsf{E}(\mathbb{1}_{\phi \geq u_n} \mathbb{1}_{M_{(j-1)(p+t)} < u_n} \circ f^{p+t-k}) \\ &\qquad - \mu(\phi \geq u_n)\mu(M_{(j-1)(p+t)} < u_n)\Big)\Bigg| \\ &\leq \sum_{k=1}^{p} 3Ce^{-\tau(p+t-k)} \leq c_0 e^{-\tau t}. \end{aligned} \tag{21}$$

Hence, by (21) and definition of $\Xi_{p,n}$, we have

$$\Gamma_j \leq t\mu(\phi \geq u_n) + 2p\Xi_{p,n} + c_0 e^{-\tau t}.$$

As in Proposition 6.1, we take $p = [l^s]$ with $s > 0$. Furthermore, we let $t = [\log(l^{\frac{2}{\tau}})]$. We therefore have

$$\Gamma_j \leq \tilde{\Gamma} := t\mu(\phi \geq u_n) + 2p\Xi_{p,n} + \frac{c_2}{l^2}.$$

for some constant c_2.

We have chosen p and l such that $p \leq l^s$ and $l \neq n$. Since $\mu(\phi \geq u_n) = O(n^{-\sigma})$, we have $1 - p\mu(\phi \geq u_n) \geq 0$ for large n, since $\sigma > \frac{1}{2} \geq s$. Proposition 6.1 now implies that

$$\mu(M_l < u_n) \leq (1 - p\mu(\phi \geq u_n))^q + \tilde{\Gamma} \sum_{j=0}^{q-1} (1 - p\mu(\phi \geq u_n))^j.$$

Using that for $0 < x < 1$

$$(1 - x)^k = \exp(k \log(1 - x)) \leq \exp(-kx),$$

we obtain that

$$\begin{aligned} \mu(M_l < u_n) &\leq \exp(-pq\mu(\phi \geq u_n)) + \tilde{\Gamma} \sum_{j=0}^{q-1} \exp(-jp\mu(\phi \geq u_n)) \\ &\leq \exp(-pq\mu(\phi \geq u_n)) + q\tilde{\Gamma}. \end{aligned}$$

Hence

$$\mu(M_l < u_n) \leq \exp(-pq\mu(\phi \geq u_n)) + qt\mu(\phi \geq u_n) + 2pq\Xi_{p,n} + \frac{c_2 q}{l^2}.$$

Since $p = [l^s]$ we have $r \leq l^s$ and $pq \geq l - l^s$. Since $\mu(\phi \geq u_n) = O(n^{-\sigma})$, we have that $l^s \mu(\phi \geq u_n)$ is bounded if s is small enough. We obtain

$$\begin{aligned}\mu(M_l < u_n) &\leq \exp(-(l - l^s)\mu(\phi \geq u_n)) + qt\mu(\phi \geq u_n) + 2pq\Xi_{p,n} + \frac{c_2}{l^{1+s}} \\ &\leq e^C \exp(-l\mu(\phi \geq u_n)) + qt\mu(\phi \geq u_n) + 2pq\Xi_{p,n} + \frac{c_2}{l^{1+s}}. \qquad \square\end{aligned}$$

So far we have not used (A2). The next corollary gives a key bound that we use in the proof of the main results. It further quantifies the e^C multiplier in equation (20), and hence gives an error bound for estimating $\mu(M_l < u_n)$ in terms of $e^{-l\mu(\phi \geq u_n)}$.

Corollary 7.3. *Let $f\colon \mathcal{X} \to \mathcal{X}$ be an interval map with an ergodic probability measure μ. We assume (A1) with rate function $\Theta(j) = \exp(-\tau j)$. Suppose the hypothesis of (A2) holds with parameters σ, γ and $s \in (0, \frac{1}{2}]$. Then there are constants $C, c_2, c_3 > 0$ such that for sufficiently large $l \leq n$ written as $l = qp + r$ with $p = [l^s]$, $0 \leq r < p$, and $t = [\log(l^{\frac{2}{\tau}})]$ we have*

$$\begin{aligned}(22) \quad |\mu(M_l < u_n) - \exp(-l\mu(\phi \geq u_n))| &\leq qt\mu(\phi \geq u_n) + 2n\Xi_{p,n} \\ &\quad + \frac{c_2}{l^{1+s}} + c_3 l^s \mu(\phi \geq u_n).\end{aligned}$$

Furthermore, there exists $\gamma' > 0$ such that for all $\beta > 0$ and $l > \beta n$

$$|\mu(M_l < u_n) - \exp(-l\mu(\phi \geq u_n))| \leq C_\beta n^{-\gamma'}. \tag{23}$$

Here, the constant $C_\beta > 0$ depends on β.

Proof. Note that we have $\sigma > \frac{1}{2}$ by assumption (A2). Then, from Proposition 6.1 and as in the proof of Lemma 7.2, we get that

$$|\mu(M_l < u_n) - \exp(-(l - l^s)\mu(\phi \geq u_n))| \leq qt\mu(\phi \geq u_n) + 2n\Xi_{p,n} + \frac{c_2}{l^{1+s}}.$$

Noting that $\mu(\phi \geq u_n) = O(n^{-\sigma})$, the constraint $\sigma > (2 + s)/3$ from (A2) implies that $l^s \mu(\phi \geq u_n)$ is bounded. Hence,

$$\begin{aligned}&|\mu(M_l < u_n) - e^{-l\mu(\phi \geq u_n)}| - |\mu(M_l < u_n) - e^{-(l-l^s)\mu(\phi \geq u_n)}| \\ &\qquad \leq |e^{-l\mu(\phi \geq u_n)} - e^{-(l-l^s)\mu(\phi \geq u_n)}| \leq |1 - e^{l^s \mu(\phi \geq u_n)}| \leq c_3 l^s \mu(\phi \geq u_n).\end{aligned}$$

This gives the equation (22) stated in the corollary. For equation (23), the existence of the constant $\gamma' > 0$ follows from equation (9). Indeed, to see this we consider each right-hand term of (22), and note that (by hypothesis) $l > \beta n$ for some $\beta > 0$, and hence $c_2 l^{-1-s}$ within (22) is $O(n^{-1-s})$. By (A2), we have $n\Xi_{p,n} < n^{-\gamma}$. Similarly we have $qt\mu(\phi \geq u_n) = O((\log n) n^{1-s-\sigma})$. The latter term is $O(n^{-\gamma_1})$ for some γ_1. This follows from the constraint $\sigma > 1 - s/2$ in (9). We have already considered the term $c_3 l^s \mu(\phi \geq u_n)$. Hence this completes the proof. $\square$

We state the following further corollary, which is an easy consequence of the results developed so far.

Corollary 7.4. *Let $f\colon \mathcal{X} \to \mathcal{X}$ be an interval map with an ergodic probability measure μ. Assume that conditions (A1) with rate function $\Theta(j) = \exp(-\tau j)$*

and (A3) hold. Suppose the hypothesis of (A2) holds with parameters σ, γ and s. Furthermore, we suppose that

$$\sum_{n=1}^{\infty} \mu(\phi \geq u_n) \exp(-\beta n \mu(\phi \geq u_n)) < \infty,$$

for some $\beta < \frac{\tau}{\log K}$, where K is the constant in Lemma 7.1. Then

$$\mu(M_n \geq u_n \, \text{ev.}) = 1.$$

We remark that the conclusion of this corollary is not optimal relative to the statements within Theorems 3.2 and 3.4. The ideas presented here will be used in the proofs of these theorems, but optimised accordingly. We also clarify within the need for the constrains imposed by equation (9) within (A2).

Proof of Corollary 7.4. Combining Lemma 7.1 and Lemma 7.2 we now get that

$$\begin{aligned}(24)\quad &\mu(M_n < u_n \text{ and } X_{n+1} > u_{n+1}) \\ &\qquad \leq e^C \exp(-l\mu(\phi \geq u_n))\mu(\phi \geq u_n) + n\Xi_{p,n}\mu(\phi > u_n) \\ &\qquad\qquad + qt\mu(\phi \geq u_n)^2 + \frac{c_2}{l^{1+s}}\mu(\phi > u_n) + c_1 K^l e^{-\tau n}.\end{aligned}$$

We take $p = n^s$, $pq \approx n$ and $l = [\beta n]$ where $\beta < \frac{\tau}{\log K}$. This makes the term $c_1 K^l e^{-\tau n}$ summable over n. Also, the term $c_2 l^{-1-s}\mu(\phi > u_n)$ is summable over n.

For the term $qt\mu(\phi \geq u_n)^2$, the relation $\sigma > 1 - s/2$ implies this is summable over n (noting that the contribution from t is $O(\log n)$). Consider the term $n\Xi_{p,n}\mu(\phi > u_n)$. By (A2), this term is summable by the assumption $\Xi_{p,n} < n^{-1-\gamma}$, and the fact that $\sigma > 1 - \gamma$. However, we still have to check a self-consistency condition involving σ and s, since we also know by exponential decay of correlations (A1) that $\Xi_{p,n} > c' n^s \mu(\phi > u_n)^2$ for some $c' > 0$. By equation (9), we have $\sigma > (2 + s)/3$, and therefore it follows that $c' n^{s+1}\mu(\phi > u_n)^3$ is also summable.

Hence, we have showed that $\mu(M_n < u_n$ and $X_{n+1} \geq u_{n+1})$ is summable provided that

$$\sum_{n=1}^{\infty} \mu(\phi \geq u_n) \exp(-\beta n \mu(\phi \geq u_n)) < \infty,$$

for some $\beta < \frac{\tau}{\log K}$, which finishes the proof. □

8. Proof of Theorem 3.2, Case (1).

The proof of Case (1) in Theorem 3.2 follows the same ideas of Section 7 that led to Corollary 7.4. The only thing missing is that Lemma 7.1 need not be true since $\mathbb{1}_{M_l < u}$ is not of bounded variation. The use of Lemma 7.1 is therefore replaced by the following lemma.

Lemma 8.1. *Under the assumptions of Theorem 3.2 there is a constant* $c_1 > 0$ *such that for every* $l, n \in \mathbb{N}$ *with* $l \leq n$ *we have*

$$\begin{aligned}\mu(M_l < u \text{ and } X_{n+1} \geq u) \\ \leq \mu(M_l < u)\mu(X_{n+1} \geq u) + c_1 l e^{-\tau(n-l)} \mu(X_{n+1} \geq u).\end{aligned}$$

Proof. We let $\mathscr{L}$ be the transfer operator

$$\mathscr{L}\psi(x) = \sum_{f(y)=x} g(y)\psi(y),$$

where $\log g$ is the potential of the Gibbs measure. The assumptions of the theorem mean the following. The density h of μ with respect to the conformal measure ν is an eigenfunction of $\mathscr{L}$ with eigenvalue $\lambda = e^P$, where P is the pressure. The eigenfunction h is of bounded variation.

The operator $\mathscr{L}$ has the following useful properties (see [44]). It satisfies $\int \psi \, \mathrm{d}\nu = \lambda^{-1} \int \mathscr{L}(\psi) \, \mathrm{d}\nu$ [44, Lemma 4.10]. If $\int \psi \, \mathrm{d}\nu = 0$, then by [44, Sub-Lemma 4.1.1] and the proof of [44, Lemma 4.11] we have

$$\lambda^{-n} \sup |\mathscr{L}^n(\psi)| \leq \lambda^{-n} \|\mathscr{L}^n(\psi)\|_{\mathrm{BV}} \leq C \|\psi\|_{\mathrm{BV}} e^{-\tau n}.$$

Using the first of these properties, we have

$$\begin{aligned}\int (\mathbb{1}_{M_l<u} - c)\mathbb{1}_{X_{n+1}\geq u} \, \mathrm{d}\mu &= \lambda^{-n} \int \mathscr{L}^n((\mathbb{1}_{M_l<u} - c)\mathbb{1}_{X_{n+1}\geq u} h) \, \mathrm{d}\nu \\ &= \lambda^{-n} \int \mathscr{L}^n((\mathbb{1}_{M_l<u} - c)h)\mathbb{1}_{X_1\geq u} \, \mathrm{d}\nu \\ &\leq \lambda^{-n} \sup |\mathscr{L}^n((\mathbb{1}_{M_l<u} - c)h)| \cdot \nu(X_1 \geq u).\end{aligned}$$

Letting $c = \mu(M_l < u)$ this implies that

$$\begin{aligned}\mu(M_l < u \text{ and } X_{n+1} \geq u) &= \int \mathbb{1}_{M_l<u}\mathbb{1}_{X_{n+1}\geq u} \, \mathrm{d}\mu \\ &\leq \mu(M_l < u)\mu(X_{n+1} \geq u) \\ &\quad + \lambda^{-n} \sup |\mathscr{L}^n((\mathbb{1}_{M_l<u} - c)h)| \cdot \nu(X_1 \geq u).\end{aligned} \tag{25}$$

It remains to estimate the above supremum. We put $\psi = (\mathbb{1}_{M_l<u} - c)h$.

Claim. *We have* $\|\mathscr{L}^l \psi\|_{\mathrm{BV}} \leq C_4 \lambda^l l$ *for some constant* C_4 *that does not depend on* l *or* u.

Proof of Claim. Clearly, we have

$$\sup |\mathscr{L}^l(\psi)| \leq \sup |\mathscr{L}^l(h)| = \lambda^l \sup |h| < \infty. \tag{26}$$

We shall now estimate $\operatorname{var} \mathscr{L}^l(\psi)$. Since $\mathbb{1}_{M_l<u} = \prod_{k=0}^{l-1} \mathbb{1}_{X_1<u} \circ f^k$, we have

$$\mathscr{L}^l(\mathbb{1}_{M_l<u} h)(x) = \sum_{f^l(y)=x} g_l(y)h(y) \prod_{k=0}^{l-1} \mathbb{1}_{X<u}(f^k y),$$

where $g_l(y) = g(y)g(f(y)) \ldots g(f^{l-1}(y))$. We let $(f^l)_j$ denote the branches of f^l (i.e. the restrictions of f^l to maximal intervals of monotonicity) and write

$$\mathscr{L}^l(\mathbb{1}_{M_l<u} h)(x) = \sum_j g_l((f^l)_j^{-1}(x))h((f^l)_j^{-1}(x)) \prod_{k=0}^{l-1} \mathbb{1}_{X_1<u}(f^k((f^l)_j^{-1}(x))).$$

Then

$$\begin{aligned}&\operatorname{var}\mathscr{L}^l(\mathbb{1}_{M_l<u}h)(x)\\&\qquad\le\sum_j\operatorname{var}\Big(g_l((f^l)_j^{-1}(x))h((f^l)_j^{-1}(x))\prod_{k=0}^{l-1}\mathbb{1}_{X<u}(f^k((f^l)_j^{-1}(x))\Big).\end{aligned}$$

Since $k<l$ we have

$$\operatorname{var}\mathbb{1}_{X_1<u}(f^k\circ(f^l)_j^{-1})\le\operatorname{var}\mathbb{1}_{X_1<u}=2.$$

Using that

$$\operatorname{var}(\phi\psi)\le\sup|\phi|\operatorname{var}\psi+\sup|\psi|\operatorname{var}\phi,$$

this implies that

$$\operatorname{var}\prod_{k=0}^{l-1}\mathbb{1}_{X_1<u}(f^k\circ(f^l)_j^{-1})\le 2l.$$

Let

$$G_j(x)=g_l((f^l)_j^{-1}(x))h((f^l)_j^{-1}(x))$$

and

$$F_j(x)=\prod_{k=0}^{l-1}\mathbb{1}_{X_1<u}(f^k((f^l)_j^{-1}(x))).$$

With this notation, we have from above that

$$\operatorname{var}\mathscr{L}^l(\mathbb{1}_{M_l<u}h)(x)\le\sum_j\operatorname{var}(G_jF_j)\le\sum_j\big(\operatorname{var}G_j\sup F_j+\sup G_j\operatorname{var}F_j\big).$$

Since

$$\lambda^l h(x)=\mathscr{L}^l(h)(x)=\sum_j G_j(x),$$

we have $\sum_j\operatorname{var}G_j=\lambda^l\operatorname{var}h$.

Hence,

$$\begin{aligned}\operatorname{var}\mathscr{L}^l(\mathbb{1}_{M_l<u}h)(x)&\le\sum_j\big(\operatorname{var}G_j+\sup G_j\operatorname{var}F_j\big)\\&\le C_0\lambda^l+C_1\lambda^l l\le C_2\lambda^l l,\end{aligned}$$

where the constant C_2 does not depend on l.

Since $\operatorname{var}\mathscr{L}^l((\mathbb{1}_{M_l<u}-c)h)\le\operatorname{var}\mathscr{L}^l(\mathbb{1}_{M_l<u}h)+c\lambda^l\operatorname{var}h\le C_3\lambda^l l$, we have now proved with the aid of (26) that

$$\|\mathscr{L}^l((\mathbb{1}_{M_l<u}-c)h)\|_{\mathrm{BV}}\le C_4\lambda^l l$$

for some constant C_4. □

We recall that the constant c was chosen so that $\int\psi\,\mathrm{d}\nu=\int(\mathbb{1}_{M_l<u}-c)h\,\mathrm{d}\nu=\int(\mathbb{1}_{M_l<u}-c)\,\mathrm{d}\mu=0$. Hence we also have $\int\mathscr{L}^l\psi\,\mathrm{d}\nu=\lambda^l\int\psi\,\mathrm{d}\nu=0$. Then we estimate

$$\begin{aligned}\lambda^{-n}\sup|\mathscr{L}^n(\psi)|&=\frac{\lambda^{-l}}{\lambda^{n-l}}\sup|\mathscr{L}^{n-l}(\mathscr{L}^l\psi)|\le\lambda^{-l}\cdot C\|\mathscr{L}^l\psi\|_{\mathrm{BV}}\cdot e^{-\tau(n-l)}\\&\le C_5le^{-\tau(n-l)}.\end{aligned}$$

It follows from (25) that

$$\mu(M_l < u \text{ and } X_{n+1} \geq u) \leq \mu(M_l < u)\mu(X_{n+1} \geq u) + C_5 l e^{-\tau(n-l)} \nu(X_1 \geq u).$$

Finally, since h is the density of μ with respect to ν, and h is bounded, we have $C_5 \nu(X_1 \geq u) \leq c_1 \mu(X_1 \geq u) = c_1 \mu(X_{n+1} \geq u)$ for some constant c_1. □

Following the proofs in Section 7 that led to Corollary 7.4, and using Lemma 8.1 instead of Lemma 7.1, we get instead of (24) that

$$\begin{aligned} \mu(M_n < u_n \text{ and } X_{n+1} > u_{n+1}) &\leq e^C \exp(-l\mu(\phi \geq u_n))\mu(\phi \geq u_n) + n\Xi_{p,n}\mu(\phi > u_n) \\ &\quad + qt\mu(\phi \geq u_n)^2 + \frac{c_4 \mu(\phi > u_n)}{l^{1+s}} + c_1 l e^{-\tau(n-l)}. \end{aligned} \tag{27}$$

We take $l = n - [n^\beta]$, where $\beta < \sigma$, and $p = n^s$, $pq \approx n$. This makes the term $c_1 l e^{-\tau(n-l)}$ as well as the term $c_4 l^{-1-s} \mu(\phi > u_n)$ in (27) summable over n. The other terms are summable as in the proof of Corollary 7.4.

The rest of the proof of Case (1) in Theorem 3.2 is the same as in the proofs outlined in Section 7, and we obtain that $\mu(M_n < u_n \text{ and } X_{n+1} \geq u_n)$ is summable provided that

$$\sum_{n=1}^{\infty} \mu(\phi \geq u_n) \exp(-n\mu(\phi \geq u_n)) < \infty.$$

This finishes the proof.

9. Proof of Theorem 3.2, Case (2)

Throughout this section, we assume that the assumptions of Theorem 3.2 hold.

To prove Theorem 3.2 we follow [20, Section 4], in particular we follow the proof of Theorem 4.3.2 within. Given $\lambda > 0$, consider the sequence $a_n := \exp\{\lambda n / \log n\}$. This choice of sequence has several properties which we elaborate on in the course of the proof. Now, for a given non-decreasing sequence (v_n), showing $\mu(M_n \leq v_n \text{ i.o.}) = 1$ can be reduced to showing $\mu(M_{b_n} \leq v_{b_n} \text{ i.o.}) > 0$ for some subsequence b_n. Since $\{M_n > u_n \text{ ev.}\} = \mathcal{H}_{\mathrm{ea}}(\mathbf{B})$, this follows from a zero–one law for eventually almost hitting sets under the assumption of ergodicity (see [36, Lemma 1]).

The following reductions are elementary manipulations, and do not depend on the precise form of (a_n), nor on the dependency structure of the process. To show $\mu(M_n \leq v_n \text{ i.o.}) = 1$, we can first reduce this to finding $c > 0$, and M_0, such that for all $M \geq M_0$ we have

$$\mu\left(\bigcup_{n=M}^{\infty} \{M_{a_n} \leq v_{a_n}\}\right) \geq c.$$

This can be reduced further to showing that for all $M > 0$, there exists $M' > M$ such that

$$\mu\left(\bigcup_{n=M}^{M'} \{M_{a_n} \leq v_{a_n}\}\right) \geq c. \tag{28}$$

Now for arbitrary events (A_n), we have

$$\mu\left(\bigcup_{n=M}^{M'} A_n\right) = \sum_{n=M}^{M'} \mu(A_n) - \sum_{n=M}^{M'} \mu\left(A_n \cap \left(\bigcup_{i=n+1}^{M'} A_i\right)\right).$$

Thus equation (28) holds if there exists $\Delta > 0$, independent of M, M' such that

$$\sum_{n=M}^{M'} \mu(M_{a_n} \leq v_{a_n}) \geq \Delta > 0,$$

and $\delta < 1$, such that for all $M_0 \leq M \leq n \leq M'$,

$$\mu\left(M_{a_n} \leq v_{a_n}, \text{ and } \bigcup_{i=n+1}^{M'} \{M_{a_i} \leq v_{a_i}\}\right) \leq \delta\mu(M_{a_n} \leq v_{a_n}).$$

Thus a requirement placed on the choice of sequence (a_n) is that

$$\sum_{n=1}^{\infty} \mu(M_{a_n} \leq v_{a_n}) = \infty, \tag{29}$$

and

$$\sum_{t=n+1}^{M'} \mu\big(\{M_{a_n} \leq v_{a_n}\} \cap \{M_{a_t} \leq v_{a_t}\}\big) \leq \delta\mu(M_{a_n} \leq v_{a_n}). \tag{30}$$

In the i.i.d. case, these conditions are shown to hold for the sequence $a_n = e^{\lambda n/\log n}$ for suitable $\lambda > 0$. The approach followed is that we can realise each term in the sum of (30) as the product

$$\mu(M_{a_n} \leq v_{a_n})\mu(M_{a_t - a_n} \leq v_{a_t}). \tag{31}$$

This uses the fact that v_{a_n} is non-decreasing. The remainder of the proof in the i.i.d. case is elementary, and uses further facts, such as

$$\mu(M_{a_n} \leq v_{a_n}) = F_X(v_{a_n})^{a_n}, \quad \mu(M_{a_t - a_n} \leq v_{a_t}) = F_X(v_{a_t})^{a_t - a_n}, \tag{32}$$

where F_X is the probability distribution function. For the dependent case, we need to recover approximate versions of (31) and (32), and show that the same proof goes through. This can be done using the mixing properties of the dynamical system, and the blocking arguments. To do this, we consider a further sequence ℓ_t, with $\ell_t < a_t - a_n$, and defined for $t > n$. Since $a_n = e^{\lambda n/\log n}$, we can choose ℓ_t to grow at various speeds, such as power law of t. The role of ℓ_t is to de-correlate successive maxima in the following way:

$$\begin{aligned}
&\mu(\{M_{a_n} \leq v_{a_n}\} \cap \{M_{a_t} \leq v_{a_t}\}) \\
&\quad = \mu(\{M_{a_n} \leq v_{a_n}\} \cap \{M_{a_t - a_n} \circ f^{a_n} \leq v_{a_t}\}) \\
&\quad \leq \mu(\{M_{a_n} \leq v_{a_n}\} \cap \{M_{a_t - a_n - \ell_t} \circ f^{a_n + \ell_t} \leq v_{a_t}\}) \\
&\quad \leq \mu(M_{a_n} \leq v_{a_n})\mu(M_{a_t - a_n - \ell_t} \leq v_{a_t}) + c_1 a_n e^{-\tau \ell_t},
\end{aligned} \tag{33}$$

where in the last line we have used a more general version of Lemma 8.1, which is proved in the same way. We choose $\ell_t = \kappa t$ for some $\kappa > 0$ to be specified in the proof below. It suffices to consider v_n such that $\mu(X_1 > v_n) \approx \log\log n/n$, with $\approx$ denoting multiplication by a constant within $[1/2, 2]$. (See [20, Lemma 4.3.2] on taking this reduction).

From the dynamical blocking arguments in Lemma 7.2 we obtain

$$\mu(\{M_{a_n} \leq v_{a_n}\}) = Ce^{-a_n\mu(X_1>v_{a_n})} + O(a_n^{-\beta}), \tag{34}$$

$$\mu(M_{a_t-a_n-\ell_t} \leq v_{a_t}) = Ce^{-(a_t-a_n-\ell_t)\mu(X_1>v_{a_t})} + O((a_t-a_n-\ell_t)^{-\beta}), \tag{35}$$

for some constants $C, \beta > 0$. Equations (34), (35) are obtained using Corollary 7.3 and equation (22) within. In particular, to obtain equation (35), we put $l = a_t - a_n - \ell_t$ in (22), and use the sequence growth estimates given in (36) (below). Our choice of ℓ_t grows fast enough to ensure decay of correlations gives a good approximation to (31), but slow enough to ensure a good approximation to (32). We state the following result.

Lemma 9.1. *Assume that* (29) *holds. Then equation* (30) *holds.*

Proof. We summarise some properties of $a_n = e^{\lambda n/\log n}$. For all $n \to \infty$, and moderate values of $t > 0$

$$\frac{a_{n+t} - a_n}{a_{n+t}} \log\log a_{n+t} \geq C\lambda t. \tag{36}$$

To see this apply the mean value theorem:

$$\begin{aligned}\frac{a_{n+t} - a_n}{a_{n+t}} &= 1 - \exp\{-\lambda(t+n)/\log(t+n) + \lambda n/\log n\} \\ &= 1 - \exp\left\{\lambda\left(-\frac{1}{\log x} + \frac{1}{(\log x)^2}\right)t\right\}, \quad (x \in [n, t+n]), \\ &\geq \frac{C\lambda t}{\log(n+t)}.\end{aligned}$$

Then note that $\log\log a_{n+t}$ is $\approx \log(n+t)$ for large n.

By assumption of (29), and given any $\Delta > 0$ we can choose $M' > M$ so that

$$\Delta \leq \sum_{n=M}^{M'} \mu(M_{a_n} \leq v_{a_n}) \leq 2\Delta. \tag{37}$$

(This is valid when $\mu(M_{a_n} \leq v_{a_n}) \to 0$, which is true in our case). Now, let us consider the right hand terms of (33). We factor out $\mu(M_{a_n} \leq v_{a_n})$ as follows,

$$\begin{aligned}&\mu(M_{a_n} \leq v_{a_n})\mu(M_{a_t-a_n-\ell_t} \leq v_{a_t}) + c_1 a_n e^{-\tau\ell_t} \\ &\qquad = \mu(M_{a_n} \leq v_{a_n})\left(\mu(M_{a_t-a_n-\ell_t} \leq v_{a_t}) + \frac{c_1 a_n e^{-\tau\ell_t}}{\mu(M_{a_n} \leq v_{a_n})}\right),\end{aligned}$$

and, hence, to show (30) it is sufficient to show the final bracketed term can be bounded by $\delta < 1$, when $\ell_t = \kappa t$, and after summing over $t \in [n+1, M']$. Consider the exponential decay of correlation term (with rate $\tau_1 = e^{-\tau} < 1$) within the bracket. This is bounded as follows.

$$c_1 a_n e^{-\tau\ell_t} \cdot \mu(M_{a_n} \leq v_{a_n})^{-1} \leq C a_n \tau_1^{\ell_t} \cdot \left(e^{-a_n\mu(X_1>v_{a_n})} + O(a_n^{-\beta})\right)^{-1}.$$

Using $\mu(X_1 > v_n) \approx \log\log n/n$, $a_n = e^{\lambda n/\log n}$ and $\ell_t = \kappa t$ gives a bound

$$C \exp\Big\{\frac{\lambda n}{\log n}\Big\} \cdot \tau_1^{\kappa t} \cdot \Big(\frac{1}{(\log a_n)^D} + O(a_n^{-\beta})\Big)^{-1}$$

with $D \in [1/2, 2]$. Now choose κ so that $e^{\lambda} < \tau_1^{-\kappa}$. This term decays exponentially fast in $t \in [n+1, M']$, and the first term (for $t = n+1$) is $o(1)$ for large n. Thus the sum of this contribution is bounded by $\delta_1 < 1$, for large n. It now suffices to consider

$$\begin{aligned}\mu(&M_{a_t - a_n - \ell_t} \le v_{a_t}) \\ &= C \exp\{-(a_t - a_n - \ell_t)\mu(X_1 > v_{a_t})\} + O((a_t - a_n - \ell_t)^{-\beta}).\end{aligned}$$

The $O(\cdot)$ term is again summable, and decays exponentially fast with $(a_t - a_n - \ell_t)^{-\beta} = o(1)$ for $t = n+1$. This is therefore bounded by $\delta_2 < 1$, for large n. We also have

$$\begin{aligned}\exp\{-(a_t &- a_n - \ell_t)\mu(X_1 > v_{a_t})\} \\ &= \exp\{-(a_t - a_n)\mu(X_1 > v_{a_t})\} \cdot \exp\{\ell_t \mu(X_1 > v_{a_t})\} \\ &= \exp\{-(a_t - a_n)\mu(X_1 > v_{a_t})\} \cdot \exp\{D\kappa t \log\log a_t / a_t\} \\ &= \exp\{-(a_t - a_n)\mu(X_1 > v_{a_t})\} \cdot e^{o(1)}, \quad (n \to \infty).\end{aligned}$$

The latter $e^{o(1)}$ comes from the precise form of a_t. Hence it suffices to show that

$$\sum_{t=n+1}^{M'} \exp\{-(a_t - a_n)\mu(X_1 > v_{a_t})\} < \delta_3,$$

for δ_3 sufficiently small. However this is now the same argument as used in [20], as it depends only on a_t, and the assumption on the asymptotics of $\mu(X_1 > v_n)$. The formalities depend on splitting $t \in [n+1, M']$ into three time windows, and the bounds utilise equations (36) and (37). □

From this lemma, we can deduce first the weaker conclusion, namely that if $\mu(X_1 > v_n) \le c \log\log n/n$ for $c < 1$, then $\mu(M_n \le v_n \text{ i.o.}) = 1$. This follows from the fact that by equation (34)

$$\mu(M_{a_n} \le v_{a_n}) = \frac{C}{(\log a_n)^c} + O(a_n^{-\beta}) \approx \frac{C}{n^c},$$

(which is not summable and, hence, (29) is satisfied), and that (30) holds for this sequence.

To complete the proof of Theorem 3.2, it is enough to show that the choice of a_n is enough to conclude that the 2nd half of the Robbins–Siegmund condition implies $\mu(M_n \le v_n \text{ i.o.}) = 1$. Following [20], we show that for a sequence v_n satisfying

$$\sum_{n=1}^{\infty} \mu(X_1 > v_n) = \infty \qquad \text{and} \qquad \sum_{n=1}^{\infty} \mu(X_1 > v_n) e^{-n\mu(X_1 > v_n)} = \infty,$$

then $\sum_{n=1}^{\infty} \mu(M_{a_n} \le v_{a_n}) = \infty$.

Since $\sum_n (a_n)^{-\beta} < \infty$, we have to show by (34) that $\sum_{n=1}^{\infty} \exp\{-a_n \mu(X_1 > v_{a_n})\} = \infty$. By monotonicity considerations, we have

$$\begin{aligned}\infty &= \sum_{n=1}^{\infty} \sum_{j=a_n}^{a_{n+1}} \mu(X_1 > v_j) e^{-j\mu(X_1 > v_j)} \\ &\leq \sum_{n=1}^{\infty} \mu(X_1 > v_{a_n})(a_{n+1} - a_n) \exp\{-a_n \mu(X_1 > v_{a_{n+1}})\}.\end{aligned}$$

Hence the implication follows as in [20].

10. Proof of Theorem 3.4

We split this section up into two parts, and treat cases (1) and (2) of Theorem 3.4 separately.

10.1. Proof of Theorem 3.4 Case (1).

Following the methods of Section 7 leading to Corollary 7.3, we obtain

$$\mu(M_l < u_n) \leq e^C \exp(-l\mu(\phi \geq u_n)) + n\Xi_{p,n} + qt\mu(\phi \geq u_n) + \frac{c_2}{l^{1+s}},$$

where $p = [l^s]$, and $t = [\log(l^{\frac{2}{\tau}})]$. Take $l = n$ and for $\rho > 0$ suppose that $\mu(\phi \geq u_n) \leq (\log n)^\rho / n$. Then by (A2) we have that

$$\mu(M_n < u_n) \leq c_6 \exp(-n\mu(\phi \geq u_n)) + O(n^{-\gamma'}) \tag{38}$$

for some $\gamma' \in (0, 1)$. We remark here that it is indeed sufficient to restrict to $\mu(\phi \geq u_n) \leq (\log n)^\rho / n$ rather than the more general case $\mu(\phi \geq u_n) \approx n^{-\sigma}$ for some $\sigma \in (0, 1)$. For the latter case the error term $n^{-\gamma'}$ in equation (38) would dominate.

Let $a > 1$ and take $n_k = [a^k]$. Then $n_k^{-\gamma'}$ is summable over k. We state the following result, which we prove at the end of this section.

Proposition 10.1. *Suppose that $n \mapsto n\mu(\phi \geq u_n)$ is non-decreasing and positive. Then for $a > 1$ and $\theta > 0$ we have*

$$\begin{aligned}\sum_{k=1}^{\infty} e^{-\theta[a^k]\mu(\phi \geq u_{[a^{k+1}]})} < \infty \qquad &\Rightarrow \qquad \sum_{n=1}^{\infty} \mu(\phi \geq u_n) e^{-n\theta\mu(\phi \geq u_n)} < \infty \\ &\Rightarrow \qquad \sum_{k=1}^{\infty} e^{-\theta[a^k]\mu(\phi \geq u_{[a^k]})} < \infty.\end{aligned}$$

Using Proposition 10.1 and (38) we have almost surely that there exists a k_0 such that $M_{n_k} \geq u_{n_k}$ for all $k \geq k_0$.

Suppose that such a k_0 exists. Let $n > n_{k_0}$, and take k such that $n_k \leq n < n_{k+1}$. Then

$$M_n \geq M_{n_k} \geq u_{n_k}.$$

Since u_n is an increasing sequence, we obtain that

$$M_n \geq u_{[n/a]}$$

holds for all $n > n_{k_0}$. This proves the first statement of the theorem.

Finally, suppose that $\mu(\phi \geq u_n) \geq c\frac{\log\log n}{n}$ for some $c > 1$. Put $\tilde{u}_n = u_{[\tilde{a}n]}$. Since $c > 1$, we can choose $\tilde{a} > a > 1$ close to one so that for large n we have $\mu(\phi \geq \tilde{u}_n) \geq \tilde{c}\frac{\log\log n}{n}$ with some $\tilde{c} > 1$ satisfying $\tilde{c}[a^k]/[a^{k+1}] > 1$.

Then

$$\sum_{k=1}^{\infty} e^{-[a^k]\mu(\phi \geq \tilde{u}_{[a^{k+1}]})} \leq \sum_{k=1}^{\infty} e^{-\tilde{c}\frac{[a^k]}{[a^{k+1}]}\log\log[a^{k+1}]} < \infty$$

holds. It follows from Proposition 10.1 that $\sum_{n=1}^{\infty} \mu(\phi \geq \tilde{u}_n) e^{-n\mu(\phi \geq \tilde{u}_n)} < \infty$. Then we can apply Case (1) of Theorem 3.4 and obtain that almost surely $M_n \geq \tilde{u}_{[n/a]}$ holds eventually. Since $1 < a < \tilde{a}$, $\tilde{u}_{[n/a]} = u_{[\tilde{a}[n/a]]} > u_n$ holds when n is large, and the result follows.

Proof of Proposition 10.1. We first prove a variant of Cauchy condensation. Suppose that $a > 1$ and that c_k is a sequence of positive numbers with $c_{k+1} \leq c_k$ for all k. Let $C = c_1 + c_2 + \ldots + c_{[a]-1}$. Then

$$\begin{aligned}\sum_{n=1}^{\infty} c_n &= C + \sum_{k=1}^{\infty} \sum_{j=[a^k]}^{[a^{k+1}]-1} c_j \leq C + \sum_{k=1}^{\infty} \sum_{j=[a^k]}^{[a^{k+1}]-1} c_{[a^k]} \\ &= C + \sum_{k=1}^{\infty} ([a^{k+1}] - [a^k]) c_{[a^k]} \leq C + \sum_{k=1}^{\infty} (a^{k+1} - a^k + 1) c_{[a^k]} \\ &\leq C + \sum_{k=1}^{\infty} a^{k+1} c_{[a^k]}.\end{aligned}$$

Hence,

$$\sum_{k=1}^{\infty} a^k c_{[a^k]} < \infty \qquad \Rightarrow \qquad \sum_{n=1}^{\infty} c_n < \infty.$$

Similarly, we have

$$\begin{aligned}\sum_{n=1}^{\infty} c_n &= C + \sum_{k=1}^{\infty} \sum_{j=[a^k]}^{[a^{k+1}]-1} c_j \geq C + \sum_{k=1}^{\infty} \sum_{j=[a^k]}^{[a^{k+1}]-1} c_{[a^{k+1}]} \\ &= C + \sum_{k=1}^{\infty} ([a^{k+1}] - [a^k]) c_{[a^{k+1}]} \geq C + \sum_{k=1}^{\infty} (a^{k+1} - 1 - a^k) c_{[a^{k+1}]} \\ &= C + \sum_{k=1}^{\infty} (a^{k+1}(1 - a^{-1}) - 1) c_{[a^{k+1}]}.\end{aligned}$$

Hence, we have proved

$$\sum_{k=1}^{\infty} a^k c_{[a^k]} < \infty \qquad \Leftrightarrow \qquad \sum_{n=1}^{\infty} c_n < \infty.$$

Now, since $n \mapsto n\mu(\phi \geq u_n)$ is non-decreasing, $n \mapsto \exp(-n\theta\mu(\phi \geq u_n))$ is non-increasing. Hence, since $[a^k]\theta\mu(X \geq u_{[a^k]}) \geq c$ for some $c > 0$, we have

$$\begin{aligned}\sum_{k=1}^{\infty} e^{-[a^k]\theta\mu(X \geq u_{[a^k]})} = \infty \ &\Rightarrow \ \sum_{k=1}^{\infty} [a^k]\theta\mu(X \geq u_{[a^k]}) e^{-[a^k]\theta\mu(X \geq u_{[a^k]})} = \infty \\ &\Leftrightarrow \ \sum_{k=1}^{\infty} \theta\mu(X \geq u_n) e^{-n\theta\mu(X \geq u_n)} = \infty.\end{aligned}$$

Finally, provided that there exists a constant $K > 0$ such that

$$[a^k]\mu(X \geq u_{[a^k]})e^{-[a^k]\theta\mu(X\geq u_{[a^k]})} \leq Ke^{-[a^{k-1}]\theta\mu(X\geq u_{[a^k]})} \tag{39}$$

for all large k, we have

$$\begin{aligned} &\sum_{n=1}^{\infty}\mu(X \geq u_n)e^{-n\theta\mu(X\geq u_n)} = \infty \\ &\qquad\Leftrightarrow \quad \sum_{k=1}^{\infty}[a^k]\mu(X \geq u_{[a^k]})e^{-[a^k]\theta\mu(X\geq u_{[a^k]})} = \infty \\ &\qquad\Rightarrow \quad \sum_{k=1}^{\infty} e^{-[a^{k-1}]\theta\mu(X\geq u_{[a^k]})} = \infty \\ &\qquad\Leftrightarrow \quad \sum_{k=1}^{\infty} e^{-[a^k]\theta\mu(X\geq u_{[a^{k+1}]})} = \infty. \end{aligned}$$

The condition (39) is implied by the condition

$$[a^k]\mu(X \geq u_{[a^k]}) \leq Ke^{(1-[a^{k-1}]/[a^k])[a^k]\theta\mu(X\geq u_{[a^k]})}.$$

Since $(1 - [a^{k-1}]/[a^k])$ is bounded away from 0 for large k, it is possible to find a K such that (39) holds for all large k. $\square$

10.2. **Proof of Theorem 3.4 Case (2).** The arguments mirror those used for proving case (2) of Theorem 3.2, but with fine adjustments used for the sequences. We start with an immediate consequence of assumptions (A1) and (A3), where the decay rate is given by $\Theta(j) = \exp(-\tau j)$. The following lemma builds upon Lemma 7.1

Lemma 10.2. *There is a constant c_1 such that for $m > n$ and $\ell \leq m - n$,*

$$\begin{aligned} &\mu(\{M_n \leq u_n\} \cap \{M_m \leq u_m\}) \\ &\qquad\qquad \leq \mu(M_n \leq u_n)\mu(M_{m-n-\ell} \leq u_m) + c_1K^n e^{-\tau\cdot(n+\ell)}. \end{aligned}$$

Proof. Let $\varphi_1(x) = \mathbb{1}_{\{M_n\leq u_n\}}(x)$ and $\varphi_2(x) = \mathbb{1}_{\{M_{m-n-\ell}\leq u_m\}}(x)$. We estimate the BV-norm of φ_1. Since for any interval A, $f^{-1}(A)$ has at most K_f connected components (by (A3)), it follows that the BV-norm of φ_1 is bounded by K^n, for some constant K.

Using decay of correlations, we get that

$$\begin{aligned} \mu(\{M_n \leq u_n\} &\cap \{M_m \leq u_m\}) \\ &= \mu(\{M_n \leq u_n\} \cap \{M_{m-n} \circ f^n \leq u_m\}) \\ &\leq \mu(\{M_n \leq u_n\} \cap \{M_{m-n-\ell} \circ f^{n+\ell} \leq u_m\}) \\ &= \int \varphi_1 \cdot \varphi_2 \circ f^{n+\ell} \,\mathrm{d}\mu \\ &\leq \int \varphi_1 \,\mathrm{d}\mu \int \varphi_2 \,\mathrm{d}\mu + Ce^{-\tau\cdot(n+\ell)}\|\varphi_1\|_{\mathrm{BV}}\|\varphi_2\|_\infty \\ &\leq \mu(M_n \leq u_n)\mu(M_{m-n-\ell} \leq u_m) + c_1K^n e^{-\tau\cdot(n+\ell)}. \end{aligned}$$ $\square$

We modify the proof of part (2) of Theorem 3.2, still following [20, Section 4]. This time we define the sequence (a_n) recursively:

$$a_{n+1} = \big(1 + (\log\log(a_n))^3\big) \cdot a_n, \qquad a_0 = \exp(\lambda)$$

for a given $\lambda > 1$.

As before we want to show that

$$\sum_{n=1}^{\infty} \mu(M_{a_n} \le v_{a_n}) = \infty, \tag{40}$$

and

$$\sum_{t=n+1}^{M'} \mu\big(\{M_{a_n} \le v_{a_n}\} \cap \{M_{a_t} \le v_{a_t}\}\big) \le \delta\mu(M_{a_n} \le v_{a_n}) \tag{41}$$

for all $M' > n$. Once again, we consider a further sequence ℓ_t, with $\ell_t < a_t - a_n$, and defined for $t > n$. Since $a_{n+1} - a_n = (\log\log(a_n))^3 \cdot a_n$, we can choose $\ell_t = (\log\log(a_{t-1})) \cdot a_{t-1}$ for all $t > n$. As before, ℓ_t is used to de-correlate successive maxima. By Lemma 10.2 we have:

$$\begin{aligned} &\mu(\{M_{a_n} \le v_{a_n}\} \cap \{M_{a_t} \le v_{a_t}\}) \\ &\qquad\qquad \le \mu(M_{a_n} \le v_{a_n})\mu(M_{a_t - a_n - \ell_t} \le v_{a_t}) + c_1 K^{a_n} e^{-\tau\cdot(a_n + \ell_t)}. \end{aligned} \tag{42}$$

By [20, Lemma 4.3.2] it suffices to consider v_n such that $\mu(X_1 > v_n) \approx \log\log n/n$, with $\approx$ denoting multiplication by a constant within $[1/2, 2]$. In addition, the dynamical blocking arguments in Lemma 7.2 (in particular using Corollary 7.3) give

$$\begin{aligned} \mu(M_{a_n} \le v_{a_n}) &= Ce^{-a_n\mu(X_1 > v_{a_n})} + O(a_n^{-\beta}), \\ \mu(M_{a_t - a_n - \ell_t} \le v_{a_t}) &= Ce^{-(a_t - a_n - \ell_t)\mu(X_1 > v_{a_t})} + O((a_t - a_n - \ell_t)^{-\beta}), \end{aligned} \tag{43}$$

for some constants $C, \beta > 0$.

Lemma 10.3. *Assume that* (40) *holds. Then equation* (41) *holds.*

Proof. By assumption of (40), and given any $\Delta > 0$ we can choose $M' > M$ so that

$$\Delta \le \sum_{n=M}^{M'} \mu(M_{a_n} \le v_{a_n}) \le 2\Delta.$$

(This is valid when $\mu(M_{a_n} \le v_{a_n}) \to 0$, which is true in our case by (43) and $\mu(X_1 > v_n) \approx \log\log n/n$). Now, let us consider the right hand terms of (42). We factor out $\mu(M_{a_n} \le v_{a_n})$ as follows,

$$\begin{aligned} &\mu(M_{a_n} \le v_{a_n})\mu(M_{a_t - a_n - \ell_t} \le v_{a_t}) + c_1 K^{a_n} e^{-\tau\cdot(a_n + \ell_t)} \\ &\qquad = \mu(M_{a_n} \le v_{a_n})\left(\mu(M_{a_t - a_n - \ell_t} \le v_{a_t}) + \frac{c_1 K^{a_n} e^{-\tau\cdot(a_n + \ell_t)}}{\mu(M_{a_n} \le v_{a_n})}\right), \end{aligned}$$

and, hence, to show (41) it is sufficient to show the final bracketed term can be bounded by $\delta < 1$, when $\ell_t = (\log\log(a_{t-1})) \cdot a_{t-1}$, and after summing over $t \in [n+1, M']$. Consider the exponential decay of correlation term within the bracket. With equation (43) this is bounded as follows.

$$\frac{c_1 K^{a_n} e^{-\tau\cdot(a_n+\ell_t)}}{\mu(M_{a_n} \le v_{a_n})} \le C K^{a_n} e^{-\tau\cdot(a_n+\ell_t)} \cdot \left(e^{-a_n \mu(X_1 > v_{a_n})} + O(a_n^{-\beta})\right)^{-1},$$

Using $\mu(X_1 > v_n) \approx \log\log n/n$ and our choices for a_n as well as ℓ_t gives a bound

$$C e^{\log(K) a_n - \tau\cdot(a_n+\ell_t)} \cdot \left(\frac{1}{(\log a_n)^D} + O(a_n^{-\beta})\right)^{-1}$$

with $D \in [1/2, 2]$. This term decays exponentially fast in $t \in [n+1, M']$, and the first term (for $t = n+1$) is $o(1)$ for large n. Thus the sum of this contribution is bounded by $\delta_1 < 1$, for large n. It now suffices to consider

$$\begin{aligned} &\mu(M_{a_t - a_n - \ell_t} \le v_{a_t}) \\ &\qquad = C \exp\{-(a_t - a_n - \ell_t)\mu(X_1 > v_{a_t})\} + O((a_t - a_n - \ell_t)^{-\beta}). \end{aligned}$$

The $O(\cdot)$ term is again summable, and decays exponentially fast with $(a_t - a_n - \ell_t)^{-\beta} = o(1)$ for $t = n+1$. This is therefore bounded by $\delta_2 < 1$, for large n. We also have

$$\begin{aligned} &\exp\{-(a_t - a_n - \ell_t)\mu(X_1 > v_{a_t})\} \\ &\qquad = \exp\{-(a_t - a_n)\mu(X_1 > v_{a_t})\} \cdot \exp\{\ell_t \mu(X_1 > v_{a_t})\} \\ &\qquad = \exp\{-(a_t - a_n)\mu(X_1 > v_{a_t})\} \cdot \exp\{D\ell_t \log\log a_t / a_t\} \\ &\qquad = \exp\{-(a_t - a_n)\mu(X_1 > v_{a_t})\} \cdot e^{o(1)}, \quad (n \to \infty). \end{aligned}$$

The latter $e^{o(1)}$ comes from the precise form of a_t and

$$\begin{aligned} D\ell_t \frac{\log\log a_t}{a_t} &= D \log\log(a_{t-1}) \frac{\log\log\left((1 + \log\log(a_{t-1}))a_{t-1}\right)}{\left(1 + (\log\log a_{t-1})^3\right)} \\ &\le D \log\log(a_{t-1}) \frac{\log\log(a_{t-1}^2)}{\left(1 + (\log\log a_{t-1})^3\right)} \\ &= D \log\log(a_{t-1}) \frac{\log 2 + \log\log(a_{t-1})}{\left(1 + (\log\log a_{t-1})^3\right)} \to 0. \end{aligned}$$

Hence it suffices to show that

$$\sum_{t=n+1}^{M'} \exp\{-(a_t - a_n)\mu(X_1 > v_{a_t})\} < \delta_3,$$

for δ_3 sufficiently small. This follows from $\mu(X_1 > v_n) \approx \log\log n/n$ and our growth of a_t. □

Now we complete the proof of case (2) in Theorem 3.4. It is enough to show that the choice of a_n is sufficient to conclude that for a sequence v_n satisfying $n \mapsto n\mu(X_1 > v_n)$ is non-decreasing,

$$\sum_{n=1}^{\infty} \mu(X_1 > v_n) = \infty, \qquad \text{and} \qquad \sum_{n=1}^{\infty} \mu(X_1 > v_n) e^{-n\gamma\mu(X_1 > v_n)} = \infty$$

for some $\gamma > 1$, then $\sum_{n=1}^{\infty} \mu(M_{a_n} \le v_{a_n}) = \infty$.

Since $\sum_n (a_n)^{-\beta} < \infty$, we have to show by (43) that

$$\sum_{n=1}^{\infty} \exp\{-a_n \mu(X_1 > v_{a_n})\} = \infty.$$

Since $n \mapsto n\mu(X_1 > v_n)$ is non-decreasing, we have by monotonicity considerations that

$$\begin{aligned}
\infty &= \sum_{n=1}^{\infty} \sum_{j=a_n}^{a_{n+1}} \mu(X_1 > v_j) e^{-j\gamma\mu(X_1 > v_j)} \\
&\leq \sum_{n=1}^{\infty} \mu(X_1 > v_{a_n})(a_{n+1} - a_n) \exp\{-a_n \gamma \mu(X_1 > v_{a_n})\} \\
&\leq \sum_{n=1}^{\infty} C \frac{\log\log(a_n)}{a_n} (a_{n+1} - a_n) \cdot e^{-(\gamma-1)D \log\log(a_n)} \cdot e^{-a_n \mu(X_1 > v_{a_n})} \\
&= \sum_{n=1}^{\infty} C (\log\log(a_n))^4 \cdot \frac{1}{(\log a_n)^{(\gamma-1)D}} \cdot \exp\{-a_n \mu(X_1 > v_{a_n})\} \\
&\leq C_1 \cdot \sum_{n=1}^{\infty} \exp\{-a_n \mu(X_1 > v_{a_n})\},
\end{aligned}$$

where we used $\gamma > 1$ in the last step. Hence, we conclude that

$$\sum_{n=1}^{\infty} \exp\{-a_n \mu(X_1 > v_{a_n})\} = \infty$$

as required. □

11. Proof of Theorems 4.1 and 4.2

To prove the theorems stated in Section 4, we need a version of equation (38) incorporating the extremal index θ. Before proving each theorem in turn, we collect relevant results from [18] which adapt the blocking methods of Section 6 to the case $\theta \in (0, 1)$.

We use the notations of Section 4, and for any $s, \ell \in \mathbb{N}$, and an event $B \subset \mathcal{X}$ we write

$$\mathcal{W}_{s,\ell}(B) := \cap_{i=s}^{s+\ell-1} f^{-i}(B^{\complement}).$$

(Notice that $\mathcal{W}_{0,n}(U(u)) = \{M_n \leq u\}$.) In the following, we recall also that q denotes the period of the hyperbolic periodic point $\tilde{x}$.

Using [18, Corollary 2.4] combined with [15, Proposition 5.1] leads to the following approximation results. Consider sequences $t_n, k_n \to \infty$, and in the following we take $\lesssim$ to mean '$\leq$' up to a uniform positive multiplying constant. Then

$$\left| \mu\big(\mathcal{W}_{0,n}(A_n^{(q)})\big) - \left(1 - \frac{n}{k_n} \mu(A_n^{(q)})\right)^{k_n} \right| \\ \lesssim k_n t_n \mu(U_n) + n\gamma_1(q, n, t_n) + n\gamma_2(q, n, k_n), \tag{44}$$

where

$$\gamma_1(q,n,t_n) = \sup_{\ell\in\mathbb{N}} |\mu(A_n^{(q)} \cap \mathcal{W}_{t_n,\ell}(A_n^{(q)})) - \mu(A_n^{(q)})\mu(\mathcal{W}_{0,\ell}(A_n^{(q)}))|,$$

$$\gamma_2(q,n,k_n) = \sum_{j=q+1}^{n/k_n} \mu(A_n^{(q)} \cap f^{-j}(A_n^{(q)})).$$

In Section 11.1 we explain how $\gamma_1(q,n,t_n)$ and $\gamma_2(q,n,k_n)$ are estimated. In addition, we have

$$\big|\mu(M_n \le u_n) - \mu\big(\mathcal{W}_{0,n}(A_n^{(q)})\big)\big| \le q\mu(U_n \setminus A_n^{(q)}). \tag{45}$$

Putting these results together leads to the following lemma.

Lemma 11.1. *Under the assumptions of equations* (44), (45) *the following formula is valid:*

$$\begin{aligned}|\mu(M_n \le u_n) - \exp(-n\theta\mu(\phi \ge u_n))| \\ \lesssim k_n t_n \mu(U_n) + n\gamma_1(q,n,t_n) + n\gamma_2(q,n,k_n) \\ + q\mu(U_n \setminus A_n^{(q)}) + |\theta_n - \theta| + \frac{n^2\mu(A_n^{(q)})^2}{k_n}.\end{aligned} \tag{46}$$

Proof. The proof requires justification of the inclusion of the last two terms. From equation (44), and using $e^x = 1 + x + O(x^2)$ we have:

$$\begin{aligned}\Big(1 - \frac{n}{k_n}\mu(A_n^{(q)})\Big)^{k_n} &= e^{-n\mu(A_n^{(q)})} + O\Big(\frac{n^2\mu(A_n^{(q)})^2}{k_n}\Big) \\ &= e^{-n\theta\mu(X_1>u_n)} + O\Big(\frac{n^2\mu(A_n^{(q)})^2}{k_n}\Big) + O(|\theta - \theta_n|). \quad \square\end{aligned}$$

11.1. **Completing the proofs of Theorem 4.1 and Theorem 4.2.** In the case of proving Theorem 4.1 we follow the proof of Theorem 3.2, while in the case of proving Theorem 4.2 we follow the proof of Theorem 3.4. In the first instance we use equation (46) to prove the following result analogous to the conclusion of Corollary 7.3.

Proposition 11.2. *Under the assumptions of Theorem 4.1 we have*

$$\mu(M_n < u_n) = \exp(-n\theta\mu(\phi \ge u_n)) + O(n^{-\gamma'})$$

for some $\gamma' \in (0,1)$.

Proof. To prove the proposition we estimate each term on the right hand side of equation (46). The term which requires careful analysis is the one involving $\gamma_2(q,n,k_n)$, and we treat that term last. First of all from (A4), we have $|\theta - \theta_n| = O(n^{-\hat{\sigma}})$. We take $t_n = c\log n$ for some $c \gg -1/\log\tau$ and apply exponential decay of correlations via (A1) to get a polynomial decay in n for $n\gamma_1(q,n,t_n)$ within equation (46). This is achieved by applying Definition 3.1 using indicator functions $\varphi_1(x) = \mathbb{1}_{A_n^{(q)}}(x)$, $\varphi_2(x) = \mathbb{1}_{\mathcal{W}_{0,\ell}(A_n^{(q)})}(x)$, and taking rate function $\Theta(t_n)$. In particular we obtain $\|\varphi_1\|_{\mathrm{BV}} < \infty$ with bound independent of n, and $\|\varphi_2\|_\infty = 1$ (thus with bound independent of ℓ).

Now, for the sequence u_n, we can restrict to the case $\mu(A_n^{(q)}) = O((\log n)^\rho/n)$ for some $\rho > 0$. The reasons are similar to the choice of u_n made in equation (38). Since we assume that $\theta \neq 0$ (again from (A4)) the term $q\mu(U_n \setminus A_n^{(q)})$ also decays to zero at the same rate $O((\log n)^\rho/n)$. Similarly the term $\frac{n^2\mu(A_n^{(q)})^2}{k_n} = O(n^{-\gamma_1})$ for some $\gamma_1 > 0$. Thus we are left to estimate the remaining term $\gamma_2(q, n, k_n)$.

By inspecting equation (46) it suffices to show existence of $B > 0$ such that

$$n \sum_{j=q+1}^{n/k_n} \mu(A_n^{(q)} \cap f^{-j}(A_n^{(q)})) = O(n^{-B}). \tag{47}$$

The constant k_n plays the same role as the blocking number p used in the proof of Theorem 3.2. We take $k_n = \sqrt{n}$, but other rates can be chosen. We split (47) into the following two sums:

$$n \sum_{j=q+1}^{n/k_n} \mu(A_n^{(q)} \cap f^{-j}(A_n^{(q)})) = \tag{48}$$
$$n \sum_{j=q+1}^{g(n)} \mu(A_n^{(q)} \cap f^{-j}(A_n^{(q)})) + n \sum_{j=g(n)}^{n/k_n} \mu(A_n^{(q)} \cap f^{-j}(A_n^{(q)})),$$

and take $g(n) = \kappa \log n$ for some $\kappa > 0$ to be determined. The following results will be useful.

Lemma 11.3. *Suppose that $\tilde{x}$ is a hyperbolic periodic point, and $|(f^q)'(\tilde{x})| \in (1, \infty)$. Then there is a time $R_n \geq c_0 \log n$ with $f^j(A_n^{(q)}) \cap A_n^{(q)} = \emptyset$ for all $j \leq R_n$.*

This is an elementary calculation based on estimating the time taken for orbits to escape from a fixed neighbourhood of the (hyperbolic) periodic orbit, see [17, Proposition 2]. The constant c_0 depends on $(f^q)'(\tilde{x})$, and on the size of the neighbourhood around $\tilde{x}$ for which f^q is a diffeomorphism.

Hence if we choose $g(n) = \kappa \log n$, with $\kappa < c_0$, then the first term on the right hand side of (48) is zero. To deal with the second term for this choice of $g(n)$, we use decay of correlations (A1) and Proposition 15.1 (see Section 15). Taking $k_n = \sqrt{n}$, we obtain that this is bounded by

$$\sum_{j=g(n)}^{\sqrt{n}} \mu(A_n^{(q)} \cap f^{-j}(A_n^{(q)})) \leq \sqrt{n}\mu(U_n)^2 + \|1_{U_n}\|_{L^{p'}} \|1_{U_n}\|_{\mathrm{BV}} \sum_{j=g(n)}^{\sqrt{n}} \Theta(j),$$
$$\leq c_1\sqrt{n}\left(\frac{(\log n)^{2\rho}}{n^2}\right) + c_2 \frac{(\log n)^{\rho/p'}}{n^{1/p'}} \cdot \frac{1}{n^{\kappa_1}},$$

with $\kappa_1 = \kappa\tau$. Hence, by choosing p' sufficiently close to 1, equation (47) holds for some $B > 0$. Therefore the conclusion of Proposition 11.2 holds. □

To complete the proof of Theorem 4.1 we now follow the proof of Theorem 3.2 step by step, as detailed in Sections 8 and 9. Similarly, in the case of proving Theorem 4.2 we follow the proof of Theorem 3.4. The inclusion of the parameter θ in the distribution for $\mu(M_n \leq u_n)$ causes no further technical obstacles in applying these methods of proof.

11.2. **Proof of Proposition 4.5.** To prove Proposition 4.5, we begin with a local analysis of the dynamics near the neutral fixed point $\tilde{x} = 0$ to estimate $\mu(A_n^{(q)})$, and then use the identities given in equations (44) and (45).

An estimate for $\mu(A_n^{(q)})$ is given in [19], and we repeat the main steps here for completeness. Take $q = 1$, and consider the ball $B(0, r_n)$, with $v_n = \psi(r_n)$. The set $A_n^{(1)}$ is precisely the set $[\hat{r}_n, r_n]$, with $f(\hat{r}_n) = r_n$. Using the fact that the density $\rho(x)$ takes the form

$$\rho(x) = \frac{H(x)}{x^a}, \quad \text{with} \quad H \in L^{1+\epsilon},$$

we obtain

$$\mu(A_n^{(1)}) \sim C_1(r_n^{1-a} - \hat{r}_n^{1-a}).$$

Using the fact that $r_n = \hat{r}_n + 2^a \hat{r}_n^{1+a}$, an asymptotic analysis yields

$$\mu(A_n^{(1)}) \sim C_2 r_n \sim C_3 \mu(X_1 > v_n)^{\frac{1}{1-a}}.$$

The constants C_i are generic constants that depend on ρ through H. Hence using equations (44) and (45), we obtain

$$\mu(M_n \leq v_n) = \exp\{-C' n \mu(X_1 > v_n)^{\frac{1}{1-a}}\} + O(n^{-\sigma'}),$$

with $C' > 0$ and $\sigma' > 0$. Using a Cauchy-condensation argument as in the proof of Theorem 3.4, we go along the sequence $n_k = b^k$ for $b > 1$. This leads to $\mu(M_{n_k} \leq v_{n_k} \text{ i.o.}) = 0$ in the case where

$$C' \mu(X_1 > v_n)^{\frac{1}{1-a}} > \frac{c \log\log n}{n},$$

for any $c > 1$. The conclusion of the proof of Proposition 4.5 follows by taking $c > (C')^{-1}$.

12. Proof of Theorem 5.1

The idea is to use the approach of Theorem 3.4. The main step is to by-pass the influence of the set $\mathcal{M}_r$. As in the proof of Theorem 3.4 we consider the sequence (r_n) such that $\mu(X_1 > \psi(r_n)) = (c\theta^{-1} \log\log n)/n$ for some $c > 1$. By the local dimension estimate at $\tilde{x}$, we have for all $\epsilon > 0$, and $r < r_0(\tilde{x})$,

$$n^{-1/d_\mu - \epsilon} < r_n < n^{-1/d_\mu + \epsilon}$$

holds for all large enough n. We now go along the subsequence (r_{n_k}) with $n_k = \lfloor a^k \rfloor$, and any $a > 1$. This leads to $\mu(\mathcal{M}_{r_{n_k}}) < Ck^{-\sigma_1}$ which is summable, and hence $\mu(\limsup_k \mathcal{M}_{r_{n_k}}) = 0$. Thus for μ-a.e. $\tilde{x} \in \mathcal{X}$ equation (17) applies along the subsequence r_{n_k} for all $k \geq k_0(\tilde{x})$. The error term is $O(k^{-\sigma_2})$, which is again summable. To complete the proof, we follow the same approach of proving Theorem 3.4, except here the error term is $O((\log n)^{-\sigma_2})$ rather than $O(n^{-\gamma})$. Along the sequence $n_k = \lfloor a^k \rfloor$, the error term remains summable. For any $c' > 1$, we therefore have $\mu(M_{n_k} \leq v_{n_k} \text{ i.o.}) = 0$, where v_n satisfies

$$\mu(X_1 > v_n) = (c'\theta^{-1} \log\log n)/n.$$

Almost surely, we have for all large k that $M_{n_k} \geq v_{n_k}$. Then, if $n \in [n_k, n_{k+1}]$, we have $M_n \geq M_{n_k} \geq v_{n_k} \geq v_{[n/a]}$, since $n \leq n_{k+1} \leq a n_k$. Hence we have that $\mu(M_n \geq v_{[n/a]} \text{ ev.}) = 1$, where $v_{[n/a]}$ is such that $\mu(X_1 > v_{[n/a]}) \sim$

$(c'a\theta^{-1}\log\log n)/n$. Since $c'>1$ is arbitrary and a can be arbitrarily chosen close to 1, we may choose c' and a such that $c'a<c$, where c is the constant related to the sequence u_n in the statement of Theorem 5.1. Then $v_{[n/a]}\geq u_n$, and this completes the proof of Theorem 5.1.

For the cases $\sigma_1<1$ or $\sigma_2<1$, then we must go along a faster growing subsequence $n_k=\lfloor e^{k^\gamma}\rfloor$ with $\gamma>1$. This is to ensure that the First Borel–Cantelli Lemma can be applied in the proof above. In particular, we must choose $\gamma>1$ so that $\gamma\sigma_1>1$ and $\gamma\sigma_2>1$. However in the window $n\in[n_k,n_{k+1}]$, the value n is not uniformly comparable to n_k. In particular we have

$$n_{k+1}/n_k\leq Ce^{(k+1)^\gamma-k^\gamma}\leq e^{ck^{\gamma-1}},$$

where c depends on γ. Almost surely in k, we have $M_n\geq M_{n_k}\geq v_{n_k}$, with $n\in[n_k,n_{k+1}]$, and

$$\mu(X_1>v_n)=(c\theta^{-1}\log\log n)/n.$$

To estimate n_k in terms of n, we have

$$n_{k+1}=\lfloor e^{(k+1)^\gamma}\rfloor\geq n\geq\lfloor e^{k^\gamma}\rfloor=n_k.$$

Noting that $n\leq n_ke^{ck^{\gamma-1}}$, and rearranging for k in terms of n, we obtain

$$n_k\geq n\exp\{-c(\log n)^{\frac{\gamma-1}{\gamma}}\}.$$

Hence $\mu(M_n\leq v_n'\text{ i.o.})=0$, where

$$\mu(X_1>v_n')\geq e^{(\log n)^{\gamma'}}n^{-1},$$

and $\gamma'>(\gamma-1)/\gamma$.

13. On Condition (A2) and its verification for selected dynamical systems

Our main arguments used to prove condition (A2) go back to Collet [12, Corollary 2.4 and Lemma 2.5], where similar estimates are proved for some non-uniformly hyperbolic maps of an interval, including quadratic maps for Benedicks–Carleson parameters. These arguments have also been carried out for other types of systems by Gupta, Holland and Nicol [24, Section 4], for instance for Lozi and Lorenz maps.

The argument starts by first estimating the measure of the set

$$\{\,x:d(x,f^jx)<r_n\text{ for some }j\leq g(n)\,\},$$

for a suitable function $g(n)$, such as $g(n)=(\log n)^\gamma$ for $\gamma>1$. The choice of $g(n)$ is chosen to grow fast enough to combat decay of correlations, i.e. so that $\Theta(g(n))\to0$ sufficiently fast. One then obtains localised estimates using the Hardy–Littlewood maximal inequality. The result is that for many systems, including quadratic maps for Benedicks–Carleson parameters and Lorenz maps, condition (A2) holds for μ a.e. point $\tilde{x}$ when μ is a measure which is absolutely continuous with respect to Lebesgue measure. Relative to the aforementioned literature, a technical aspect in our case is that we need to assume a wider class of sequences r_n to check (A2), in particular allowing for $\mu(B(\tilde{x},r_n))\approx n^{-\sigma}$ for $\sigma<1$. In the usual extreme value theory literature, the sequences r_n are chosen so that $n\mu(B(\tilde{x},r_n))\to\ell\in(0,\infty)$, such as in equation (4) (see [45]).

To verify (A2), we give the argument in detail for the following systems: piecewise differentiable maps satisfying assumptions on decay of correlations, and piecewise expanding maps with an invariant measure which is absolutely continuous with respect to Lebesgues measure. We also explain how (A2) is obtained for quadratic maps with Benedicks–Carleson parameters, contrasting to the methods given by Collet [12].

13.1. **Condition (A2) for piecewise differentiable maps.** We consider an interval map $f\colon \mathcal{X} \to \mathcal{X}$ preserving an ergodic measure μ which is piecewise differentiable. That is, we assume that the derivative of f is uniformly bounded, so that there is a constant $\Lambda_+ \in \mathbb{R}$ with $|f'(x)| < \Lambda_+$ for all $x \in \mathcal{X}$. We allow for f to have a finite number of discontinuities, and we let $\mathcal{S}$ denote the finite set of discontinuity points. We also assume the following regularity condition on the measure μ: there exist c_1, $c_2 > 0$ and $s_1 > s_2 > 0$ such that for μ-a.e. $x \in \mathcal{X}$, there exists $r_0 > 0$ such that for all $r < r_0$,

$$c_1 r^{s_1} \leq \mu(B(x,r)) \leq c_2 r^{s_2}. \tag{49}$$

Furthermore, we assume that the upper bound holds for all $x \in \mathcal{X}$ so that r_0 is independent of x relative to the constants c_2, s_2. Examples include beta-transformations, $x \mapsto \beta x \mod 1$ $(\beta > 1)$, and the quadratic map for Benedicks–Carleson parameters. For these maps it is known that the density of μ is in L^p for $p < 2$, [49]. Hence the upper bound of (49) holds for some $s_2, c_2, r_0 > 0$ and all $x \in \mathcal{X}$, $r < r_0$. We have the following proposition.

Proposition 13.1. *Suppose that $f\colon \mathcal{X} \to \mathcal{X}$ is a piecewise differentiable interval map, preserving an ergodic measure μ. Suppose that (A1) holds, and μ satisfies equation* (49)*. Then for μ-a.e. $\tilde{x} \in \mathcal{X}$ condition (A2) holds. That is, for μ-a.e. $\tilde{x} \in \mathcal{X}$, there exists γ, s and $\sigma > 0$ such that equations* (9) *and* (10) *hold for all sequences r_n with $\mu(B(\tilde{x}, r_n)) = O(n^{-\sigma})$.*

Remark 13.2. The main conclusion of Proposition 13.1 is that (A2) applies to μ-a.e. $\tilde{x} \in \mathcal{X}$. It is possible to check (A2) point-wise under knowledge of recurrence properties of $\tilde{x}$, such as knowing that $\tilde{x}$ is pre-periodic to a hyperbolic fixed point. In these cases it is possible to remove some of the global assumptions, such as requiring existence of $\Lambda_+ < \infty$, or requiring uniformity of the constants in (49) to all $x \in \mathcal{X}$.

Remark 13.3. The proof we give is much shorter relative to the (general) methods presented in [12]. The main trick is that for piecewise differentiable systems it is sufficient to control the recurrence of typical points over a time window of order $\log n$. Previous methods have taken a longer time window of the order $(\log n)^\gamma$ for some $\gamma > 1$.

Proof. To prove this result, consider for $p_n = n^s$ and $s > 0$ the quantity $\Xi_{p_n,n}(r_n)$. Introducing an intermediate growing function $g(n) = o(p_n)$ we split up $\Xi_{p_n,n}(r_n)$ into two sums as follows:

$$\Xi_{p_n,n}(r_n) = \sum_{j=1}^{g(n)} \mu(f^j B(\tilde{x}, r_n) \cap B(\tilde{x}, r_n)) + \sum_{j=g(n)+1}^{n^s} \mu(f^j B(\tilde{x}, r_n) \cap B(\tilde{x}, r_n)). \tag{50}$$

For $\sigma > (0,1)$ we assume that $\mu(B(\tilde{x}, r_n)) = O(n^{-\sigma})$. For $g(n) = \kappa \log n$, the first sum on the right-hand side of (50) is zero by the following claim.

Claim. *There exists $\kappa > 0$, such that for μ-a.e. $\tilde{x}$, all sufficiently large n, and all $j \leq \kappa \log n$, we have $f^j B(\tilde{x}, r_n) \cap B(\tilde{x}, r_n) = \emptyset$.*

Proof of Claim. We consider the set of closely returning points $E_{r,n}$ defined by

$$E_{r,n} = \{\, x : d(f^n(x), x) < r \,\}.$$

Using Lemma 15.2 (see Appendix), condition (A1), and the regularity condition (49) we deduce that

$$\mu(E_{r,n}) \leq \int \mu(B(x, r))\, \mathrm{d}\mu(x) + De^{-\eta n} \leq c_2 r^{s_2} + De^{-\eta n},$$

for some $\eta > 0$. Let $r = 2^{-j}$, then by the regularity condition (49), and the First Borel–Cantelli Lemma we have $\mu(\liminf E^{\complement}_{r_j,j}) = 1$. Hence for μ-almost all x, there exists $j_0(x)$ such that $\mathrm{dist}(f^j(x), x) > 2^{-j}$ for all $j \geq j_0(x)$. Take $\tilde{x}$ to be a representative in this full measure set.

We impose a further restriction on the orbit of $\tilde{x}$ as follows. Let

$$F_j = \{\, x : \mathrm{dist}(f^j(x), \mathcal{S}) < 2^{-j} \,\},$$

where $\mathcal{S}$ denotes the discontinuities of f. Then we take $\tilde{x} \in \liminf F_j^{\complement}$. Again, this set also has μ-measure 1.

By (49) and the assumption $\mu(B(\tilde{x}, r_n)) = O(n^{-\sigma})$, it follows that $r_n = O(n^{-\frac{\sigma}{s_1}})$. Take $\tilde{x}$ to be in the set of μ-measure 1 as described above. We consider a time $R \equiv R(\tilde{x}, r_n)$ such that

(i) f^j is continuous on $B(\tilde{x}, r_n)$ for all $j \leq R$.
(ii) $f^j B(\tilde{x}, r_n) \cap B(\tilde{x}, r_n) = \emptyset$ for all $j \leq R$.

We provide a lower bound for R such that the above two assumptions are satisfied. First, there is a time $j_0(\tilde{x})$ for which simultaneously $d(f^j(\tilde{x}), \tilde{x}) > 2^{-j}$, and $d(f^j(\tilde{x}), \mathcal{S}) > 2^{-j}$ hold for all $j \geq j_0(\tilde{x})$. The set of such $\tilde{x}$ has μ-measure one. We let $m_0(\tilde{x})$ denote

$$m_0(\tilde{x}) = \inf_{y \in B(\tilde{x}, r_n)} \{\, d(f^j(y), \tilde{x}),\, d(f^j(y), \mathcal{S}) : j \leq j_0(\tilde{x}) \,\}.$$

For all n sufficiently large, we have $m_0(\tilde{x}) > r_n$ (perhaps removing a further countable set of $\tilde{x}$ that meet $\mathcal{S}$ before time $j_0(\tilde{x})$).

To bound R, we claim that there exists $\kappa_1 > 0$ with $R > \kappa_1 \log n$. Let $y \in B(\tilde{x}, r_n)$, then for $j \geq j_0(\tilde{x})$

$$\begin{aligned} d(f^j(y), \mathcal{S}) &\geq d(f^j(\tilde{x}), \mathcal{S}) - d(f^j(y), f^j(\tilde{x})) \\ &\geq 2^{-j} - 2r_n\Lambda_+^j, \end{aligned} \tag{51}$$

where Λ_+ is the upper bound for $|f'|$. Now for (i) to hold, we require

$$2^{-j} - 2r_n\Lambda_+^j > 0,$$

for all $j \leq R$, otherwise the ball meets the singularity set prior to time R. Hence, if we choose

$$R < \frac{-\log(r_n) - \log 2}{\log 2 + \log \Lambda_+},$$

then (i) is satisfied.

To verify item (ii), a similar bound is obtained. Similarly to above, we have for $y \in B(\tilde{x}, r_n)$ that for $j \geq j_0(\tilde{x})$

$$\begin{aligned} d(f^j(y), \tilde{x}) &\geq d(f^j(\tilde{x}), \tilde{x}) - d(f^j(y), f^j(\tilde{x})) \\ &\geq 2^{-j} - 2r_n\Lambda_+^j. \end{aligned}$$

We require that $2^{-j} - 2r_n|\Lambda_+|^j > r_n$ for all $j \leq R$. Hence, if we choose

$$R < \frac{-\log r_n - \log 3}{\log 2 + \log \Lambda_+},$$

then (ii) is satisfied.

Take any $\kappa < \frac{\sigma}{2s_1 \log(2\Lambda_+)}$ and let $g(n) = \kappa \log n$. From the above two requirements together with $r_n = O(n^{-\frac{\sigma}{s_1}})$, it follows that we may take $R = \kappa \log n$ for large n.

Note that κ depends on σ, but without loss we can further restrict to $\sigma > 1/2$ so that σ is bounded away from zero. It is immediate from the construction above that $f^j B(\tilde{x}, r_n) \cap B(\tilde{x}, r_n) = \emptyset$ for all $j \leq \kappa \log n$. □

Hence in the estimate for $\Xi_{p_n,n}$, the first sum on the right of (50) is zero. For the second sum, we use exponential decay of correlations for BV against $L^\infty(\mu)$ in conjunction with Proposition 15.1. This gives

$$\sum_{j=g(n)+1}^{n^s} \mu(f^j B(\tilde{x}, r_n) \cap B(\tilde{x}, r_n)) \leq n^s \mu(B(\tilde{x}, r_n))^2 + C_1 \mu(B(\tilde{x}, r_n))^{1/p'} e^{-\tau g(n)},$$

with $C_1 > 0$. By choice of r_n the right is bounded by:

$$n^{s-2\sigma} + n^{-\sigma/p'} \cdot n^{-\kappa_2}, \tag{52}$$

where κ_2 depends on κ and τ. Hence, there exists a choice of constants s, σ consistent with (9) so that (52) is bounded by $n^{-1-\gamma}$ for some $\gamma > 0$. This completes the proof. □

13.2. **Condition (A2) for piecewise expanding maps.** In this section we consider piecewise expanding maps. Relative to Section 13.1 we allow for unbounded derivative. This allows us to cover the Gauß map. The set up is as follows. Suppose that $f\colon [0,1] \to [0,1]$ is a piecewise expanding map, with finitely many pieces of continuity. There is then a partition $\mathscr{P} = \{I_1, \dots, I_m\}$ such that f is differentiable on each I_k. Let $\mathscr{P}_n$ be the corresponding partition for f^n. Since the partition $\mathscr{P}$ is finite, there is a $\delta_0 > 0$ such that every partition element of $\mathscr{P}$ has a diameter of at least δ_0. We let S be the set of endpoints of partition elements of $\mathscr{P}$. The set S is δ_0 separated.

Alternatively, we assume that the partition $\mathscr{P}$ is countable, in which case we assume that there is a $\delta_0 > 0$ such that for all n holds $|f^n(I)| \geq \delta_0$ whenever $I \in \mathscr{P}_n$.

We assume that f is uniformly expanding, i.e. that there is a constant Λ such that $|f'| \geq \Lambda$. Moreover, we assume that f has bounded distortion, and that μ is an ergodic measure with exponential decay of correlations for functions of bounded variation against L^1. This means that there exists a constant C such that

$$x, y \in I \in \mathscr{P}_n \qquad \Rightarrow \qquad C^{-1} \leq \frac{Df^n(x)}{Df^n(y)} \leq C$$

and

$$\left| \int \phi_1 \cdot \phi_2 \circ f^j \,\mathrm{d}\mu - \int \phi_1 \,\mathrm{d}\mu \int \phi_2 \,\mathrm{d}\mu \right| \leq C e^{-\tau j} \|\phi_1\|_{\mathrm{BV}} \|\phi_2\|_1$$

for some $\tau > 0$.

We will prove that for any such piecewise expanding map, the set of points $\tilde{x}$ which satisfies assumption (A2) has full measure. Example of a systems satisfying our assumption are piecewise expanding maps with finitely many pieces and an absolutely continuous invariant measure μ; the Gauß map with the Gauß measure; or the first return map to $[\frac{1}{2}, 1)$ for a Manneville–Pomeau map with an absolutely continuous invariant measure μ.

Proposition 13.4. *Let f be an interval map satisfying the assumptions stated above. Suppose that there is a constant c such that $\mu(I) \leq c|I|$ for any interval I. Then for μ-a.e. $\tilde{x} \in \mathcal{X}$, there exists γ, s and $\sigma > 0$ such that equations* (9) *and* (10) *hold for all non-increasing sequences (r_n) with $\mu(B(\tilde{x}, r_n)) = O(n^{-\sigma})$, and satisfying the additional assumption: for any $t > 0$,*

$$\limsup_{k\to\infty} \frac{r_{\lfloor k^t \rfloor}}{r_{\lfloor (k+1)^t \rfloor}} < \infty. \tag{53}$$

Remark 13.5. Requirement of assumption (53) is a consequence of the method of proof. Unlike in the proof of Proposition 13.1, we cannot infer that $f^j B(\tilde{x}, r_n) \cap B(\tilde{x}, r_n) = \emptyset$ for all $j = O(\log n)$. However it is possible to check (A2) point-wise if certain recurrence properties of $\tilde{x}$ are known (such as pre-periodic), and in turn relax assumption (53).

To prove Proposition 13.4, we will need two lemmata. For the first lemma, we let $A_n(\delta) = \{\, I \in \mathscr{P}_n : |f^n(I)| < \delta \,\}$, and as always, $\cup A_n(\delta) = \cup_{I \in A_n(\delta)} I$.

Lemma 13.6. *If μ satisfies $\mu(I) \leq c|I|$ for any interval I, then there exists a constant K_0 such that*

$$\mu(\cup A_n(\delta)) \leq K_0\delta$$

holds for any $\delta < \delta_0$.

Proof. We only have to consider the case when f is piecewise expanding with finitely many pieces, since in the case with countably many pieces, our assumptions imply that $\mu(\cup A_n(\delta)) = 0$ whenever $\delta < \delta_0$.

Let $I_n(x)$ denote the partition element of $\mathscr{P}_n$ which x belongs to. If $|I_n(x)| < \delta < \delta_0$ then there are $j, k < n$ such that $j \neq k$ and both $f^j(x)$ and $f^k(x)$ are close to S. More precisely, we must have

$$d(f^j(x), S) < \delta\Lambda^{-n+j} \qquad \text{and} \qquad d(f^k(x), S) < \delta\Lambda^{-n+k},$$

since otherwise, S would not have "cut" the partition element $I_n(x)$ in a way such that $|I_n(x)| < \delta$. We therefore have $\cup A_n(\delta) \subset B_n(\delta)$, where $B_n(\delta)$ is defined by

$$B_n(\delta) = \bigcup_{0 \leq j < k < n} \big(f^{-j} S_{(\delta\Lambda^{-n+j})} \cap f^{-k} S_{(\delta\Lambda^{-n+k})}\big),$$

and $S_{(\varepsilon)} = \{\, t \in [0,1] : d(t, S) < \varepsilon \,\}$. We shall estimate the measure of $B_n(\delta)$.

By decay of correlations, we have for $j < k$ that

$$\begin{aligned}
\mu\big(f^{-j} S_{(\delta\Lambda^{-n+j})} \cap f^{-k} S_{(\delta\Lambda^{-n+k})}\big) &= \mu\big(S_{(\delta\Lambda^{-n+j})} \cap f^{-(k-j)} S_{(\delta\Lambda^{-n+k})}\big) \\
&\leq \mu(S_{(\delta\Lambda^{-n+j})})\mu(S_{(\delta\Lambda^{-n+k})}) + C2(m+2)\mu(S_{(\delta\Lambda^{-n+k})})e^{-\tau(k-j)} \\
&\leq c^2\delta^2\Lambda^{-2n+j+k} + cC2(m+2)\delta\Lambda^{-n+k}e^{-\tau(k-j)}.
\end{aligned}$$

We obtain that

$$\mu(B_n(\delta)) \leq \sum_{0 \leq j < k < n} \Big(c^2\delta^2\Lambda^{-2n+j+k} + cC2(m+2)\delta\Lambda^{-n+k}e^{-\tau(k-j)}\Big) \leq K_0\delta,$$

for some constant K_0. □

We now consider the set

$$E_{j,r} = \{\, x : d(x, f^j x) < 2r \,\}.$$

In the arguments that follow, we need to control the measure of this set in terms of r when j is small. Thus we cannot use directly Lemma 15.2.

Lemma 13.7. *If μ satisfies $\mu(I) \leq c|I|$ for any interval I, then there exists a constant K_1 such that*

$$\mu(E_{j,r}) \leq K_0\delta + K_1 r\delta^{-1}$$

holds for any $4r < \delta < \delta_0$. In particular, there is a constant K_2 such that

$$\mu(E_{j,r}) \leq K_2\sqrt{r},$$

when $4r < \delta_0^2$.

Proof. Suppose that $4r < \delta < \delta_0$. By bounded distortion, we have for any $I \in \mathscr{P}_j$ with $|f^j(I)| \geq \delta$ that $I \cap E_{j,r}$ is an interval of length at most $4Cr\delta^{-1}|I|$. Together with Lemma 13.6, we get that

$$\mu(E_{j,r}) \leq \mu(\cup A_j(\delta)) + K_1 r\delta^{-1} \leq K_0\delta + K_1 r\delta^{-1}.$$

When $4r < \delta_0^2$, we may choose $\delta = \sqrt{r}$ to obtain

$$\mu(E_{j,r}) \leq (K_0 + K_1)\sqrt{r}. \qquad \square$$

We are now in position to prove Proposition 13.4.

Proof of Proposition 13.4. Let $\tilde{\gamma} > 1$, and put

$$E_k = \bigcup_{j=1}^{2(\log k)^{\tilde{\gamma}}} E_{j,r_k}.$$

Then

$$\mu(E_k) \leq K_2 (\log k)^{\tilde{\gamma}} \sqrt{r_k}, \tag{54}$$

for some constant K_2, by Lemma 13.7.

Put

$$g(x) = \sup_{r>0} \frac{1}{2r} \int_{B(x,r)} \mathbb{1}_{E_k} \, \mathrm{d}\mu = \sup_{r>0} \frac{1}{2r} \int_{B(x,r)} \mathbb{1}_{E_k}(t) h(t) \, \mathrm{d}t,$$

where h is the density of μ. By the Hardy–Littlewood maximal inequality applied to the function $\mathbb{1}_{E_k} h$, the set

$$F_k(c) = \{ x : g(x) > c \}$$

has Lebesgue measure at most $\frac{3}{c} \int \mathbb{1}_{E_k}(t) h(t) \, \mathrm{d}t = \frac{3}{c} \mu(E_k)$. Hence

$$\mu(F_k(c)) \leq \frac{3C}{c} K_2 (\log k)^{\tilde{\gamma}} \sqrt{r_k}.$$

Note that

$$x \in F_k(c)^{\complement} \qquad \Rightarrow \qquad \mu(E_k \cap B(x, r_k)) \leq 2c r_k.$$

For constants $\alpha, \beta > 0$, let $n_k = k^\beta$ and $c = n_k^{-\alpha}$. We obtain

$$\mu(F_{n_k}(n_k^{-\alpha})) \leq 3CK_2 k^{\alpha\beta} r_{n_k}^{\frac{1}{2}} (\beta \log k)^{\tilde{\gamma}}.$$

Assuming that $r_k = O(k^{-\sigma})$ for some $\sigma > 0$, we have

$$\mu(F_{n_k}(n_k^{-\alpha})) \leq 3CK_2 k^{\frac{\beta}{2}(2\alpha - \sigma)} (\beta \log k)^{\tilde{\gamma}}.$$

Take $0 < 2\alpha < \sigma$, β large enough that

$$\sum_k \mu(F_{n_k}(n_k^{-\alpha})) < \infty.$$

Hence $\mu(\limsup_{k\to\infty} F_{n_k}(n_k^{-\alpha})) = 0$ and we have for a.e. $\tilde{x}$ that

$$\mu(E_{n_k} \cap B(\tilde{x}, r_{n_k})) \leq 2 n_k^{-\alpha} r_{n_k}$$

holds for all large k (depending on $\tilde{x}$). Let such an $\tilde{x}$ be fixed. For piecewise expanding maps we can assume a stronger form of equation (49), namely we assume that $\tilde{x}$ is such that there exists a constant $c_0(\tilde{x}) > 0$ such that

$$c_0^{-1} r < \mu(B(\tilde{x}, r)) < c_0 r \tag{55}$$

holds for all $0 < r < 1$, since this is a property which holds for a.e. $\tilde{x}$. We then have for a.e. $\tilde{x}$ that

$$\mu(E_{n_k} \cap B(\tilde{x}, r_{n_k})) \leq 2c_0 n_k^{-\alpha} \mu(B(\tilde{x}, r_{n_k}))$$

holds for all large k.

Consider any large $n > 0$ and take k such that $n_k \leq n < n_{k+1}$. We then have

$$\tilde{E}_n = \bigcup_{j=1}^{(\log n)^{\tilde{\gamma}}} E_{j,r_n} \subseteq E_{n_k}$$

when k is large. Since r_n is a non-increasing sequence, we also have $B(\tilde{x}, r_n) \subseteq B(\tilde{x}, r_{n_k})$. Hence

$$\begin{aligned} \mu(B(\tilde{x}, r_n) \cap \tilde{E}_n) &\leq \mu(B(\tilde{x}, r_{n_k}) \cap E_{n_k}) \\ &\leq 2c_0 n_k^{-\alpha} \mu(B(\tilde{x}, r_{n_k})) \\ &\leq 4c_0 n^{-\alpha} \mu(B(\tilde{x}, r_{n_k})), \end{aligned}$$

if n and k are large. By (53) and (55), it follows that there exists a constant K such that

$$\mu(B(\tilde{x}, r_n) \cap \tilde{E}_n) \leq K n^{-\alpha} \mu(B(\tilde{x}, r_n))$$

holds for all n.

Suppose that $x \in B(\tilde{x}, r_n) \cap f^{-j} B(\tilde{x}, r_n)$ for some $j \leq (\log n)^{\tilde{\gamma}}$. Then $d(x, f^j x) < 2r_n$ and hence $x \in B(\tilde{x}, r_n) \cap \tilde{E}_n$. We therefore have $B(\tilde{x}, r_n) \cap f^{-j} B(\tilde{x}, r_n) \subset B(\tilde{x}, r_n) \cap \tilde{E}_n$ and

$$\mu(B(\tilde{x}, r_n) \cap f^{-j} B(\tilde{x}, r_n)) \leq \mu(B(\tilde{x}, r_n) \cap \tilde{E}_n) \leq K n^{-\alpha} \mu(B(\tilde{x}, r_n)).$$

To complete the proof, it suffices to estimate $\Xi_{p_n,n}$. We can split as in equation (50), but this time take $g(n) = (\log n)^{\tilde{\gamma}}$. The arguments above show that the first right-hand term of (50) is $O(n^{-1-\gamma})$ for a choice σ consistent with equation (9). Similarly using condition (A1), the second right-hand term of (50) is also $O(n^{-1-\gamma})$, again for a choice of constants consistent with (9). □

13.3. **Further remarks on Condition (A2) for quadratic maps.** We consider $f = f_a \colon [0,1] \to [0,1]$ defined by $f_a(x) = ax(1-x)$. For some parameters, including the parameters described by Benedicks and Carleson, there is an f_a-invariant probability measure μ_a which is equivalent with respect to Lebesgue measure. When a is a Benedicks–Carelson parameter, (f_a, μ_a) has exponential decay of correlations for functions of bounded variation against L^1 as proved by Young [49]. As remarked upon in Section 13.1, Proposition 13.1 applies to this family of maps. It is also possible to apply the methods used in the proof of Proposition 13.4. However this requires imposing the regularity condition (53) on the sequence r_n. Indeed, under dynamical assumptions that capture the quadratic map, Collet proved [12, Corollary 2.4] that there exists a constant $\beta' \in (0,1)$ such that the set

$$\tilde{E}_k = \{\, x : d(x, f_a^j(x)) < k^{-1} \text{ for some } j \leq (\log k)^5 \,\}$$

satisfies

$$\mu(\tilde{E}_k) \leq C k^{-\beta'}.$$

Let

$$E_k := \{\, x : d(x, f_a^j(x)) < r_k \text{ for some } j \leq (\log k)^4 \,\}.$$

Since $\mu(B(\tilde{x}, r_n)) = O(n^{-\sigma})$, we can use the regularity conditions (49), (53) for μ and the sequence r_n respectively to deduce that $E_k \subset \tilde{E}_{k^b}$ for some $b > 0$. Hence

$$\mu(E_k) \leq Ck^{-\beta_0},$$

for some $\beta_0 > 0$. Replacing (54) by the above estimate in the proof of Proposition 13.4 allows us to deduce that condition (A2) applies. (Within, let $\alpha \in (0, \beta_0)$ and take β sufficiently large.)

14. Appendix A — The blocking argument

We follow [12] to prove the blocking argument.

14.1. **Assumptions.**

We consider a dynamical system $(\mathcal{X}, f, \mu)$ where $\mathcal{X}$ is an interval and μ is a probability measure. In this section, we will prove Proposition 6.1. To do so, we only need to assume that μ is invariant, but when using Proposition 6.1 it shall be necessary to assume mixing.

14.1.1. *Notation.* We have an observable $\phi\colon \mathcal{X} \to \mathbb{R}$. Let $X_k = \phi \circ f^{k-1}$ and $M_n = \max\{X_1, \ldots, X_n\}$.

14.2. **Preparations.**

Lemma 14.1 (Collet [12, Proposition 3.2]). *Let t, r, m, k, p be non-negative integers. Then*

$$0 \leq \mu(M_r < u) - \mu(M_{r+k} < u) \leq k\mu(\phi \geq u) \tag{56}$$

and

$$\left|\mu(M_{m+p+t} < u) - \mu(M_m < u) + \sum_{j=1}^{p} \mathsf{E}(\mathbb{1}_{\phi \geq u} \mathbb{1}_{M_m < u} \circ f^{p+t-j})\right| \leq t\mu(\phi \geq u) + 2p \sum_{j=1}^{p} \mathsf{E}(\mathbb{1}_{\phi \geq u} \mathbb{1}_{\phi \geq u} \circ f^j). \tag{57}$$

Proof. We have $\{M_r < u\} \supset \{M_{r+k} < u\}$ and

$$\{M_r < u\} \setminus \{M_{r+k} < u\} \subseteq \bigcup_{j=r+1}^{r+k} \{X_j \geq u\}.$$

Hence

$$0 \leq \mu(M_r < u) - \mu(M_{r+k} < u) \leq \sum_{j=r+1}^{r+k} \mu(X_j \geq u) = k\mu(\phi \geq u),$$

which is (56).

We have

$$\mathbb{1}_{M_{m+p+t} < u} = \mathbb{1}_{M_p < u} \mathbb{1}_{M_t < u} \circ f^p \mathbb{1}_{M_m < u} \circ f^{p+t}.$$

Therefore,

$$\begin{aligned}0 &\leq \mathbb{1}_{M_p<u}\mathbb{1}_{M_m<u} \circ f^{p+t} - \mathbb{1}_{M_{m+p+t}<u} \\ &= \mathbb{1}_{M_p<u}\mathbb{1}_{M_m<u} \circ f^{p+t} - \mathbb{1}_{M_p<u}\mathbb{1}_{M_t<u} \circ f^{p}\mathbb{1}_{M_m<u} \circ f^{p+t} \\ &= \mathbb{1}_{M_p<u}\mathbb{1}_{M_m<u} \circ f^{p+t}(1 - \mathbb{1}_{M_t<u} \circ f^p) \\ &\leq 1 - \mathbb{1}_{M_t<u} \circ f^p = \mathbb{1}_{M_t\geq u} \circ f^p.\end{aligned}$$

It then follows that

$$\begin{aligned}&\left|\mathsf{E}\,\mathbb{1}_{M_{m+p+t}<u} - \mathsf{E}(\mathbb{1}_{M_p<u}\mathbb{1}_{M_m<u} \circ f^{p+t})\right| \\ &\qquad \leq \mathsf{E}(\mathbb{1}_{M_t\geq u} \circ f^p) = \mu(M_t \geq u) \\ &\qquad = \mu\Big(\bigcup_{k=1}^{t}\{\phi \circ f^k \geq u\}\Big) \leq t\mu(\phi \geq u). \qquad (58)\end{aligned}$$

Since

$$\begin{aligned}&\{M_m \circ f^{p+t} < u\} \setminus \{M_m \circ f^{p+t} < u \text{ and } M_p < u\} \\ &\qquad = \bigcup_{k=1}^{p}\{X_k \geq u\} \cap \{M_m \circ f^{p+t} < u\},\end{aligned}$$

we have

$$\mathsf{E}\,\mathbb{1}_{M_m<u} - \sum_{k=1}^{p}\mathsf{E}(\mathbb{1}_{X_k\geq u}\mathbb{1}_{M_m<u} \circ f^{p+t}) \leq \mathsf{E}(\mathbb{1}_{M_p<u}\mathbb{1}_{M_m<u} \circ f^{p+t}).$$

By the inclusion–exclusion inequality, we also have

$$\begin{aligned}&\mathsf{E}(\mathbb{1}_{M_p<u}\mathbb{1}_{M_m<u} \circ f^{p+t}) \\ &\qquad \leq \mathsf{E}\,\mathbb{1}_{M_m<u} - \sum_{k=1}^{p}\mathsf{E}(\mathbb{1}_{X_k\geq u}\mathbb{1}_{M_m<u} \circ f^{p+t}) \\ &\qquad\qquad + \sum_{k=1}^{p}\sum_{\substack{l=1\\ l\neq k}}^{p}\mathsf{E}(\mathbb{1}_{X_k\geq u}\mathbb{1}_{X_l\geq u}\mathbb{1}_{M_m<u} \circ f^{p+t}).\end{aligned}$$

It follows that

$$\begin{aligned}&\left|\mathsf{E}(\mathbb{1}_{M_p<u}\mathbb{1}_{M_m<u} \circ f^{p+t}) - \mathsf{E}\,\mathbb{1}_{M_m<u} + \sum_{k=1}^{p}\mathsf{E}(\mathbb{1}_{X_k\geq u}\mathbb{1}_{M_m<u} \circ f^{p+t})\right| \\ &\qquad \leq \sum_{k=1}^{p}\sum_{\substack{l=1\\ l\neq k}}^{p}\mathsf{E}(\mathbb{1}_{X_k\geq u}\mathbb{1}_{X_l\geq u}\mathbb{1}_{M_m<u} \circ f^{p+t}) \\ &\qquad \leq \sum_{k=1}^{p}\sum_{\substack{l=1\\ l\neq k}}^{p}\mathsf{E}(\mathbb{1}_{X_k\geq u}\mathbb{1}_{X_l\geq u}) \leq 2p\sum_{k=1}^{p}\mathsf{E}(\mathbb{1}_{\phi\geq u}\mathbb{1}_{X_k\geq u}). \qquad (59)\end{aligned}$$

The estimates (58) and (59) together with the triangle inequality imply (57). □

14.3. **Proof of Proposition 6.1.** We will now prove Proposition 6.1.

Let l be a large number and $s \in (0, \frac{1}{2}]$. Put $p = [l^s]$ and write l as $l = pq+r$ where $0 \leq r < p$.

We have by (56) that

$$\mu(M_{pq} < u) - \mu(M_{q(p+t)} < u) \leq qt\mu(\phi \geq u).$$

If $r \leq qt$, then $l = pq + r \leq q(p+t)$ and $\mu(M_l < u) - \mu(M_{q(p+t)} < u) \geq 0$. However, we have $q \sim l^{1-s}$ and $t \geq 1$, so $r < p \leq q \leq qt$ holds for all large enough l, since $1 - s \geq s$. Hence, when l is large, we have

$$\begin{aligned} 0 &\leq \mu(M_l < u) - \mu(M_{q(p+t)} < u) \\ &\leq \mu(M_{pq} < u) - \mu(M_{q(p+t)} < u) \leq qt\mu(\phi \geq u), \end{aligned}$$

and

$$|\mu(M_l < u) - \mu(M_{q(p+t)} < u)| \leq qt\mu(\phi \geq u).$$

Let

$$\Sigma_j = \sum_{k=1}^{p} \mathsf{E}(\mathbb{1}_{\phi \geq u} \mathbb{1}_{M_{(j-1)(p+t)} < u} \circ f^{p+t-k}).$$

By the triangle inequality we have

$$\begin{aligned} &|\mu(M_{j(p+t)} < u) - (1 - p\mu(\phi \geq u))\mu(M_{(j-1)(p+t)} < u)| \\ &\qquad \leq \left|p\mu(\phi \geq u)\mu(M_{(j-1)(p+t)} < u) - \Sigma_j\right| \\ &\qquad\qquad + \left|\mu(M_{j(p+t)} < u) - \mu(M_{(j-1)(p+t)} < u) + \Sigma_j\right|, \end{aligned}$$

and (57) with $m = (j-1)(p+t)$ implies that

$$\begin{aligned} &|\mu(M_{j(p+t)} < u) - (1 - p\mu(\phi \geq u))\mu(M_{(j-1)(p+t)} < u)| \\ &\qquad \leq \Gamma_j := \left|p\mu(\phi \geq u)\mu(M_{(j-1)(p+t)} < u) - \Sigma_j\right| \\ &\qquad\qquad + t\mu(\phi \geq u) + 2p \sum_{k=1}^{p} \mathsf{E}(\mathbb{1}_{\phi \geq u} \mathbb{1}_{\phi \geq u} \circ f^k). \end{aligned} \tag{60}$$

Let $\eta = 1 - p\mu(\phi \geq u)$. Now, using (60) iteratively, we get

$$\begin{aligned} &|\mu(M_{q(p+t)} < u) - \eta^q| \\ &\qquad \leq |\mu(M_{q(p+t)} < u) - \eta\mu(M_{(q-1)(p+t)} < u)| \\ &\qquad\quad + |\eta\mu(M_{(q-1)(p+t)} < u) - |\eta|^q| \\ &\qquad \leq \Gamma_q + |\eta||\mu(M_{(q-1)(p+t)} < u) - \eta^{q-1}| \\ &\qquad \ldots \\ &\qquad \leq \Gamma_q + |\eta|\Gamma_{q-1} + \ldots + |\eta|^{q-1}\Gamma_1. \end{aligned}$$

This proves Proposition 6.1.

15. Appendix B — On correlation decay and recurrence.

In this section we collect some useful results on decay of correlation estimates, and recurrence time distributions. In particular, these results are used for checking condition (A2). These results might also have broader interest.

15.1. **Decay of correlation estimates.** In this section we explain how condition (A1) can be improved to having (exponential) decay of correlations for $\mathcal{B}_1 = \mathrm{BV}$ versus $\mathcal{B}_2 = L^p$, (with $p > 1$).

The set up is a interval map $f\colon \mathcal{X} \to \mathcal{X}$ with an invariant probability measure μ. For $\varphi : \mathcal{X} \to \mathbb{R}$, recall the L^p norms for $p \in [1, \infty]$ is defined by

$$\|\varphi\|_p = \left(\int |\varphi|^p \,\mathrm{d}\mu \right)^{\frac{1}{p}}$$

for $p < \infty$, and $\|\varphi\|_\infty = \sup |\varphi|$. The bounded variation norm $\|\varphi\|_{\mathrm{BV}} = \|\varphi\|_\infty + \operatorname{var} \varphi$, where $\operatorname{var} \varphi$ is the total variation of φ on $\mathcal{X}$. We have the following result.

Proposition 15.1. *Let $f\colon \mathcal{X} \to \mathcal{X}$ be an interval map with an invariant probability measure μ. Suppose that correlations decay exponentially for* BV *versus L^∞. For any $p > 1$, correlations decay exponentially for* BV *versus L^p.*

Our proof relies on the Banach–Steinhaus theorem.

Proof. Fix $p > 1$ and suppose that $\varphi \in L^p$ and $\psi \in \mathrm{BV}$. (Within this section, ψ will denote such a BV function: it is not to be confused with the observable used in previous sections.) We note that $\varphi \in L^q$ for any $q \le p$, that $\|\varphi\|_q \le \|\varphi\|_p$ for such q, and that

$$\mu\big(\{\, x : |\varphi(x)| \ge t \,\}\big) \le \frac{1}{t^q} \|\varphi\|_q^q \tag{61}$$

for any $t > 0$.

Take a positive number m, which will be chosen more precisely later. We write φ as a sum $\varphi = \varphi_1 + \varphi_2$, where φ_1 is defined by

$$\varphi_1 = \mathbb{1}_{\{\, x : |\varphi(x)| \le m \,\}} \varphi.$$

Then $\|\varphi_1\|_\infty \le m$ and $\varphi_2 \in L^p$ with support in the set $\{\, x : |\varphi(x)| \ge m \,\}$. We shall first estimate the L^q norm of φ_2 for $q < p$.

Take $q < p$ and let $r, s > 1$ be such that $\frac{1}{r} + \frac{1}{s} = 1$ and $sq \le p$. We have by Hölder's inequality and (61) that

$$\|\varphi_2\|_q \le \left(\int_{\{\, x : |\varphi(x)| \ge m \,\}} |\varphi|^q \,\mathrm{d}\mu \right)^{\frac{1}{q}} = \left(\int \mathbb{1}_{\{\, x : |\varphi(x)| \ge m \,\}} |\varphi|^q \,\mathrm{d}\mu \right)^{\frac{1}{q}}$$

$$\le (\mu(\{\, x : |\varphi(x)| \ge m \,\}))^{\frac{1}{rq}} \|\varphi\|_{sq} \le \frac{1}{m^{\frac{p}{rq}}} \|\varphi\|_p^{\frac{p}{rq}} \|\varphi\|_{sq} \le \frac{1}{m^{\frac{p}{rq}}} \|\varphi\|_p^{1 + \frac{p}{rq}}. \tag{62}$$

We now consider the correlation between φ and ψ. Let

$$C(\varphi, \psi, n) = \left| \int \varphi \circ f^n \psi \,\mathrm{d}\mu - \int \varphi \,\mathrm{d}\mu \int \psi \,\mathrm{d}\mu \right|.$$

By the decomposition $\varphi = \varphi_1 + \varphi_2$ and the triangle inequality, we have

$$C(\varphi, \psi, n) \le C(\varphi_1, \psi, n) + C(\varphi_2, \psi, n).$$

Using the decay of correlations for BV against L^∞, we get

$$C(\varphi_1, \psi, n) \le C e^{-\tau n} \|\varphi_1\|_\infty \|\psi\|_{\mathrm{BV}} \le C e^{-\tau n} m \|\psi\|_{\mathrm{BV}}.$$

The correlation with φ_2 is estimated using the triangle inequality and (62) with $q = 1$ and $s = p$. We get

$$\begin{aligned} C(\varphi_2, \psi, n) &\leq \left| \int \varphi_2 \circ f^n \psi \, \mathrm{d}\mu \right| + \left| \int \varphi_2 \, \mathrm{d}\mu \int \psi \, \mathrm{d}\mu \right| \\ &\leq 2\|\varphi_2\|_1 \|\psi\|_\infty \leq \frac{2}{m^{\frac{p}{r}}} \|\varphi\|_p^{1+\frac{p}{r}} \|\psi\|_{\mathrm{BV}} = \frac{2}{m^{p-1}} \|\varphi\|_p^p \|\psi\|_{\mathrm{BV}}. \end{aligned}$$

Combining these estimates, we get

$$C(\varphi, \psi, n) \leq \Big(Ce^{-\tau n} m + \frac{2}{m^{p-1}} \|\varphi\|_p^p \Big) \|\psi\|_{\mathrm{BV}}.$$

Choose $m = e^{\frac{\tau}{p} n}$. Then

$$C(\varphi, \psi, n) \leq (C + 2\|\varphi\|_p^p) e^{-(1-\frac{1}{p})\tau n} \|\psi\|_{\mathrm{BV}}.$$

In particular, for any $\varphi \in L^p$ and $\psi \in \mathrm{BV}$ there is a constant $c(\varphi, \psi)$ such that

$$C(\varphi, \psi, n) \leq c(\varphi, \psi) e^{-(1-\frac{1}{p})\tau n}.$$

Now, an argument by Collet [11], using the Banach–Steinhaus theorem, implies that there is a constant c such that for any $\varphi \in L^p$ and $\psi \in \mathrm{BV}$ holds

$$C(\varphi, \psi, n) \leq c e^{-(1-\frac{1}{p})\tau n} \|\varphi\|_p \|\psi\|_{\mathrm{BV}}.$$

Hence (f, μ) has exponential decay of correlations for L^p against BV. □

15.2. **Estimates on recurrence time statistics.** A key argument in checking condition (A2) is understanding the distribution of recurrent points in the sense of finding the measure of the set:

$$E_{r,n} = \{\, x : d(f^n(x), x) < r \,\},$$

in terms of n and r. We have the following result.

Lemma 15.2. *Suppose that* $([0,1], f, \mu)$ *has exponential decay of correlations for* L^∞ *against* BV*, that is*

$$\left| \int \phi \circ f^n \psi \, \mathrm{d}\mu - \int \phi \, \mathrm{d}\mu \int \psi \, \mathrm{d}\mu \right| \leq C \|\phi\|_\infty \|\psi\|_{BV} e^{-\tau n}.$$

Assume that μ *satisfies* $\mu(B(x,r)) \leq cr^s$ *for some constants* $c, s > 0$ *and any ball* $B(x,r)$*.*

Then there exists a constant D *and a number* $\eta \in (0, \tau)$ *such that for any* $r > 0$ *and*

$$E_{r,n} = \{\, x : d(f^n(x), x) < r \,\}$$

we have

$$\begin{aligned} \mu(E_{r,n}) &\leq \int \mu(B(x,r)) \, \mathrm{d}\mu(x) + De^{-\eta n} \\ &\leq cr^s + De^{-\eta n}. \end{aligned}$$

Remark 15.3. This result builds upon those stated within [37, Section 4].

Proof. Let $\{I_k\}$ be a partition of $[0,1]$ into $e^{\frac{\tau}{2}n}$ intervals of equal length. Let y_k be the mid point of I_k. Put $\delta = \frac{1}{2}e^{-\frac{\tau}{2}n}$.

The function

$$F(x,y) = \begin{cases} 1 & \text{if } d(x,y) < r \\ 0 & \text{otherwise} \end{cases}$$

is such that $\mu(E_{r,n}) = \int F(f^n(x), x)\,\mathrm{d}\mu(x)$. We approximate F by $\tilde{F}$ defined by

$$\tilde{F}(x,y) = \sum_k J_k(x) \mathbb{1}_{I_k}(y),$$

where $J_k = \mathbb{1}_{(y_k - r - \delta, y_k + r + \delta)}$. Then $F \leq \tilde{F}$ holds and

$$\sum_k \int J_k \,\mathrm{d}\mu \int \mathbb{1}_{I_k} \,\mathrm{d}\mu = \iint \tilde{F} \,\mathrm{d}\mu \mathrm{d}\mu.$$

Using decay of correlations we get

$$\begin{aligned} \mu(E_{r,n}) &= \int F(f^n(x), x)\,\mathrm{d}\mu(x) \\ &\leq \int \tilde{F}(f^n(x), x)\,\mathrm{d}\mu(x) \\ &= \sum_k \int J_k(f^n(x)) \mathbb{1}_{I_k}(x)\,\mathrm{d}\mu(x) \\ &\leq \sum_k \Bigl(\int J_k \,\mathrm{d}\mu \int \mathbb{1}_{I_k} \,\mathrm{d}\mu + 3Ce^{-\tau n} \Bigr). \end{aligned}$$

Since the sum contains $e^{\frac{\tau}{2}n}$ terms, we obtain

$$\mu(E_{r,n}) \leq \iint \tilde{F} \,\mathrm{d}\mu \mathrm{d}\mu + 3Ce^{-\frac{\tau}{2}n}.$$

Finally, if we let

$$G(x,y) = \begin{cases} 1 & \text{if } d(x,y) < r + \delta \\ 0 & \text{otherwise} \end{cases}$$

then $\tilde{F} \leq G$ and

$$\begin{aligned} \iint \tilde{F} \,\mathrm{d}\mu \mathrm{d}\mu \leq \iint G \,\mathrm{d}\mu \mathrm{d}\mu &= \int \mu(B(x, r+\delta))\,\mathrm{d}\mu(x) \\ &\leq \int \mu(B(x,r))\,\mathrm{d}\mu(x) + 2c\delta^s \\ &= \int \mu(B(x,r))\,\mathrm{d}\mu(x) + 2^{1-s}ce^{-\frac{s\tau}{2}n}. \end{aligned}$$

This proves the lemma with $D = 3C + 2^{1-s}c$ and $\eta = \min(\frac{\tau}{2}, \frac{s\tau}{2})$. □

References

[1] J. S. Athreya, *Logarithm laws and shrinking target properties*, Proc. Indian Acad. Sci. Math. Sci. 119 (2009), no. 4, 541–557.

[2] D. Azevedo, A. Freitas, J. Freitas and F. Rodrigues. *Extreme Value Laws for Dynamical Systems with Countable Extremal Sets.* J. Stat. Phys, **167**, (2007), 1244–1261.

[3] O. Barndorff-Neilson, *On the rate of growth of the partial maxima of a sequence of independent identically distributed random variables.* Math. Scand. **9**, (1961), 383–394.

[4] L. Barreira, Ya. Pesin and J. Schmeling, *Dimension and product structure of hyperbolic measures*, Ann. of Math. 149 (1999), no. 3, 755–783.

[5] M. Benedicks and L.-S. Young, *Markov extensions and decay of correlations for certain Hénon maps*, Asterisque No. 261 (2000), xi, 13–56.

[6] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, **27**, of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1989.

[7] M. Carney, M. P. Holland and M. Nicol. *Extremes and extremal indices for level set observables on hyperbolic systems.* Nonlinearity, **34**, (2021) no. 2.

[8] M. Carney, M. Nicol and H. K. Zhang, *Compound Poisson law for hitting times to periodic orbits in two-dimensional hyperbolic systems*, J. Stat. Phys. 169 (2017), 804–823.

[9] J. Chazzottes and P. Collet, *Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems 33 (2013), no. 1, 49–80.

[10] N. Chernov, D. Kleinbock, *Dynamical Borel-Cantelli lemmas for Gibbs measures*, Isr. J. Math. 122 (2001), no. 1, 1–27.

[11] P. Collet, *A remark about uniform de-correlation prefactors*, unpublished note.

[12] P. Collet, *Statistics of closest return for some non-uniformly hyperbolic systems*, Ergodic Theory Dynam. Systems, **21** (2001), 401–420.

[13] M. Carvalho, A. C. M. Freitas, J. M. Freitas, M. Holland and M. Nicol, *Extremal dichotomy for hyperbolic toral automorphisms*, Dyn. Syst. 30 (2015), no. 4, 383–403.

[14] P. Embrechts, C. Klüpperlberg, and T. Mikosch, *Modelling extremal events. For insurance and finance*, Applications of Mathematics 33, Springer-Verlag, Berlin, 1997, ISBN: 3-540-60931-8.

[15] J. Freitas, *Extremal behaviour of chaotic dynamics*, Dyn. Syst. 28 (2013), no. 3, 302–332.

[16] J. Freitas, A. Freitas and M. Todd, *Hitting times and extreme value theory*, Probab. Theory Related Fields **147(3)**, 675–710, 2010.

[17] A. C. M. Freitas, J. M. Freitas, M. Todd, *The extremal index, hitting time statistics and periodicity*, Adv. Math. 231, no. 5, 2012, 2626–2665.

[18] A. C. M. Freitas, J. M. Freitas, M. Todd, *Speed of convergence for laws of rare events and escape rates*, Stochastic Process. Appl. 125 (2015), no. 4, 1653–1687.

[19] A. C. M. Freitas, J. M. Freitas, M. Todd and S. Vaienti, *Rare events of the Manneville–Pomeau map*, Stochastic Process. Appl. 126 (2016), no. 11, 3463–3479.

[20] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, John Wiley and Sons, 1978.

[21] S. Galatolo, M. P. Holland, T. Persson and Y. Zhang, *Birkhoff sums of infinite observables and anomalous time-scaling of extreme events in infinite systems*, Discrete and Continuous Dynamical Systems 41 (2021), no. 4, 1799–1841.

[22] Ch. Ganotaki, T. Persson, *On eventually always hitting points*, arXiv:2010.07714.

[23] A. Ghosh, M. Kirsebom, and P. Roy, *Continued fractions, the Chen-Stein method and extreme value theory*, Ergodic Theory Dynam. Systems 41, (2), (2021), 461–470.

[24] C. Gupta, M. P. Holland and M. Nicol, *Extreme value theory for dispersing billiards, Lozi maps and Lorenz maps*, Ergodic Theory Dynam. Systems 31, (5), (2011), 1363–1390.

[25] C. Gupta, M. Nicol and W. Ott, *A Borel-Cantelli lemma for non-uniformly expanding dynamical systems*, Nonlinearity 23, (8), (2010), 1991–2008.

[26] N. Haydn, M. Nicol, T. Persson and S. Vaienti, *A note on Borel–Cantelli lemmas for non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems 33, no. 2, (2013), 475–498.

[27] M. P. Holland and M. Nicol, *Speed of convergence to an extreme value distribution for non-uniformly hyperbolic dynamical systems*, Stochastics and Dynamics, 15, No. 4 (2015).

[28] M. P. Holland, M. Nicol and A. Török, *Extreme value distributions for non-uniformly expanding dynamical systems*, Trans. Amer. Math. Soc., 364 (2012), 661–688.

[29] M. P. Holland, M. Nicol and A. Török, *Almost sure convergence of maxima for chaotic dynamical systems*, Stochastic Process. Appl. 126 (2016), no. 10, 3145–3170.

[30] M. P. Holland, P. Rabassa and A. E. Sterk, *Quantitative recurrence statistics and convergence to an extreme value distribution for non-uniformly hyperbolic dynamical systems*, Nonlinearity 29 (2016), no. 8, 2355–2394.

[31] N. Haydn and K. Wasilewska, *Limiting distribution for error terms for the number of visits to balls in non-uniformly hyperbolic dynamical systems*, Discrete Contin. Dyn. Syst. 36 (2016), no. 5, 2585–2611.

[32] D. Kelmer, *Shrinking targets for discrete time flows on hyperbolic manifolds*, Geom. Funct. Anal. 27 (2017), no. 5, 1257–1287.

[33] D. Kelmer and H. Oh, *Exponential mixing and shrinking targets for geodesic flow on geometrically finite hyperbolic manifolds*, arXiv:1812.05251.

[34] D. Kelmer and S. Yu, *Shrinking targets problems for flows on homogeneous spaces*, Trans. Amer. Math. Soc. 372 (2019), 6283–6314.

[35] D. Kim, *The dynamical Borel–Cantelli lemma for interval maps*, Discrete Contin. Dyn. Syst. 17 (2007), no. 4, 891–900.

[36] M. Kirsebom, P. Kunde, and T. Persson, *Shrinking targets and eventually always hitting points for interval maps*, Nonlinearity 33 (2020), no. 2, 892–914.

[37] M. Kirsebom, P. Kunde, and T. Persson, *On shrinking targets and self-returning points*, arXiv:2003.01361.

[38] M. Klass, *The minimal growth rate of partial maxima*, Ann. Prob. 12, (1984), 380–389.

[39] M. Klass, *The Robbins–Siegmund series criterion for partial maxima*, Ann. Prob. 13, (4), (1985), 1369–1370.

[40] D. Kleinbock, I. Konstantoulas and F. K. Richter, *Zero–one laws for eventually always hitting points in mixing systems*, arXiv:1904.08584.

[41] D. Kleinbock, N. Wadleigh, *An inhomogeneous Dirichlet Theorem via shrinking targets*, Compositio Mathematica, 155 (2019), no. 7, 1402–1423.

[42] A. Lasota, J. A. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc. 186 (1973), 481–488 (1974).

[43] Leadbetter, M. R., Lindgren, G., and Rootzen, H. (1983), Extremes and Related Properties of Random Sequences and Processes. Springer-Verlag, New York.

[44] C. Liverani, B. Saussol and S. Vaienti, *Conformal measure and decay of correlation for covering weighted systems*, Ergod. Theory Dynam. Systems 18(6) (1998), 1399–1420.

[45] V. Lucarini et al., *Extremes and Recurrence in Dynamical Systems*, Pure and Applied Mathematics (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2016, ISBN: 978-1-118-63219-2.

[46] F. Pène and B. Saussol, *Poisson law and some nonuniformly hyperbolic systems with polynomial rate of mixing*, Ergodic Theory Dynam. Systems 36 (2016), no. 8, 2602–2626.

[47] W. Philipp, *A conjecture of Erdös on continued fractions*, Acta Arithmetica 28 (1976), Issue: 4, 379–386.

[48] M. Rychlik, *Bounded variation and invariant measures*, Studia Mathematica 76 (1983), 69–80.

[49] L.-S. Young, *Decay of Correlations for Certain Quadratic Maps*, Communications in Mathematical Physics 146 (1992), 123–138.

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

Mark Holland, University of Exeter, Department of Mathematics, North Park Road, Exeter EX4 4QF, UK
Email address: m.p.holland@exeter.ac.uk

Maxim Kirsebom, University of Hamburg, Department of Mathematics, Bundesstrasse 55, 20146 Hamburg, Germany
Email address: maxim.kirsebom@uni-hamburg.de

Philipp Kunde, Faculty of Mathematics and Computer Science, Jagiellonian University in Kraków, ul. Lojasiewicza 6, 30-348 Kraków, Poland.
Email address: pkunde.math@gmail.com

Tomas Persson, Centre for Mathematical Sciences, Lund University, Box 118, 221 00 Lund, Sweden
Email address: tomasp@maths.lth.se