A BOREL-CANTELLI LEMMA FOR NONUNIFORMLY EXPANDING DYNAMICAL SYSTEMS

CHINMAYA GUPTA, MATTHEW NICOL, AND WILLIAM OTT

Abstract. Let $(A_n)_{n=1}^{\infty}$ be a sequence of sets in a probability space $(X, \mathcal{B}, \mu)$ such that $\sum_{n=1}^{\infty} \mu(A_n) = \infty$. The classical Borel-Cantelli lemma states that if the sets A_n are independent, then $\mu(\{x \in X : x \in A_n \text{ for infinitely many values of } n\}) = 1$. We present analogous dynamical Borel-Cantelli lemmas for certain sequences of sets (A_n) in X (including nested balls) for a class of deterministic dynamical systems $T : X \to X$ with invariant probability measures. Our results apply to a class of Gibbs-Markov maps and 1-dimensional nonuniformly expanding systems modeled by Young towers. We discuss some applications of our results to the extreme value theory of deterministic dynamical systems.

1. Introduction

Borel-Cantelli lemmas are a fundamental tool used to establish the almost-sure behavior of random variables. For example, a Borel-Cantelli lemma is used in the standard proof that Brownian motion has a version with continuous sample paths. In this paper we establish dynamical Borel-Cantelli lemmas for 1-dimensional $(1D)$ nonuniformly expanding maps and give some applications of these results to the extreme value theory of dynamical systems.

Suppose $(X, \mathcal{B}, \mu)$ is a probability space. For a measurable set $A \subset X$, let $\mathbf{1}_A$ denote the characteristic function of A. We abbreviate the standard probability theory terms "infinitely often" by i.o., "almost every" by a.e., and "almost surely" by a.s.. The phrases a.e. and a.s. have the same meaning; when using them we choose the phrase that seems stylistically more natural in a given statement. The classical Borel-Cantelli lemmas (see for example [11, Section 4]) state that

(1) if $(A_n)_{n=0}^{\infty}$ is a sequence of sets in $\mathcal{B}$ and $\sum_{n=0}^{\infty} \mu(A_n) < \infty$ then $\mu(\{x \in X : x \in A_n \text{ i.o.}\}) = 0$;
(2) if $(A_n)_{n=0}^{\infty}$ is a sequence of independent events in $\mathcal{B}$ and $\sum_{n=0}^{\infty} \mu(A_n) = \infty$, then

$$\frac{\sum_{i=0}^{n-1} \mathbf{1}_{A_i}}{\sum_{i=0}^{n-1} \mu(A_i)} \to 1 \quad \text{a.s..}$$

Note that (1) does not require independence. In the dynamical setting suppose $T : X \to X$ is a measure-preserving transformation of the probability space $(X, \mathcal{B}, \mu)$ and (A_n) is a sequence of sets such that $\sum_n \mu(A_n) = \infty$. We are interested in the following question: does $T^n(x) \in A_n$ occur for infinitely many values of n for μ a.e. $x \in X$ and, if so, is there a quantitative estimate of the asymptotic number of entry times? For example, the sequence (A_n) may be a nested sequence of intervals, a setting which is often called the shrinking target problem. The assumption of independence of the events $T^{-n} A_n$ is seldom valid for deterministic dynamical systems; thus establishing Borel-Cantelli lemmas is a more difficult task. In this paper we establish results analogous to (1) and (2) for certain classes of nested intervals in the setting of 1-dimensional nonuniformly expanding dynamical systems: Theorems 2 and 3. To do this, we establish a more general Borel-Cantelli lemma for sequences of intervals in Gibbs-Markov systems: Theorem 1.

First we establish some notation. Suppose $T : X \to X$ is a measure-preserving transformation of the probability space $(X, \mathcal{B}, \mu)$. Suppose that $(A_n)_{n=0}^{\infty}$ is a sequence of sets in $\mathcal{B}$ such that $\sum_{n=0}^{\infty} \mu(A_n) = \infty$.

Date: June 21, 2010.

2000 *Mathematics Subject Classification.* Primary: 37A50; Secondary: 60G70.

Key words and phrases. Borel-Cantelli, extreme value theory, Young tower.

The research of MN was supported in part by the National Science Foundation grants DMS-0607345 and DMS-0600927. MN would also like to thank the Mittag-Leffler Institute, Djursholm, Sweden for hospitality and support during the writing of this paper.

For $n \in \mathbb{N}$, let $E_n = \sum_{i=0}^{n-1} \mu(A_i)$ and define $S_n : X \to \mathbb{Z}^+$ by

$$S_n(x) = \sum_{i=0}^{n-1} \mathbf{1}_{A_i} \circ T^i(x).$$

Definition 1.1. We will call the sequence (A_n) a

(1) ***Borel-Cantelli (BC) sequence*** if $\mu(\{x \in X : T^n(x) \in A_n \text{ i.o.}\}) = 1$;

(2) ***strong Borel-Cantelli (sBC) sequence*** if

$$\lim_{n\to\infty} \frac{S_n(x)}{E_n} = 1 \quad \text{a.s.};$$

(3) ***dense Borel-Cantelli (dBC) sequence*** with respect to the measure γ if there exists $C > 0$ for which

$$\varliminf_{n\to\infty} \frac{S_n(x)}{\sum_{i=0}^{n-1} \gamma(A_i)} \geqslant C \quad \text{a.s.}.$$

Remark 1.2. An example of a sequence that is dBC with respect to Lebesgue measure λ but not sBC with respect to the invariant measure μ is given by a sequence of nested balls (A_n) of radius $\frac{1}{n}$ about a point p in a 1-dimensional system $T : X \to X$ in which the invariant measure μ is absolutely continuous with respect to Lebesgue measure but $(d\mu/d\lambda)(p) \neq 1$ and (A_n) is sBC with respect to μ. In this setting we could take $C = (d\mu/d\lambda)(p)$. Kim [21] shows that this phenomena exists for certain intermittent maps of the interval (see also the discussion below).

There have been some results on Borel-Cantelli lemmas for uniformly hyperbolic systems. Chernov and Kleinbock [8] establish the sBC property for certain families of cylinders in the setting of topological Markov chains and for certain classes of dynamically-defined rectangles in the setting of Anosov diffeomorphisms preserving Gibbs measures. Dolgopyat [10] has related results for sequences of balls in uniformly partially hyperbolic systems preserving a measure equivalent to Lebesgue.

More recently, Kim [21] has established the sBC property for sequences of intervals in the setting of 1-dimensional piecewise-expanding maps f with $1/|f'|$ of bounded variation. Gibbs-Markov maps of the interval do not necessarily have the property that $1/|f'|$ is of bounded variation (see Remark 2.5).

Kim uses this result to prove some sBC results for nonuniformly expanding maps with an indifferent fixed point. More precisely, he considers intermittent maps of the form

$$T_\alpha(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & \text{if } 0 \leqslant x < \frac{1}{2}; \\ 2x - 1 & \text{if } \frac{1}{2} \leqslant x \leqslant 1. \end{cases} \tag{1}$$

These maps are sometimes called Liverani-Saussol-Vaienti maps [24]. If $0 < \alpha < 1$ then T_α admits an invariant probability measure μ that is absolutely continuous with respect to Lebesgue measure λ. Kim shows that if (I_n) is a sequence of intervals in $(d, 1]$ for some $d > 0$ and $\sum_n \mu(I_n) = \infty$ then I_n is an sBC sequence if (a) $I_{n+1} \subset I_n$ for all n (nested intervals) or (b) $\alpha < (3 - \sqrt{2})/2$. Kim shows that the condition $I_n \subset (d, 1]$ for some $d > 0$ is in some sense optimal (with respect to the invariant measure μ) by showing that setting $A_n = [0, n^{-1/(1-\alpha)})$ gives a sequence such that $\sum_n \mu(A_n) = \infty$ yet the sBC property does not hold; in fact, $T_\alpha^n(x) \in A_n$ for only finitely many values of n for μ a.e. $x \in [0, 1]$.

For the same class of maps T_α, Gouëzel [15] considers Lebesgue measure λ (rather than μ) and shows that if (I_n) is a sequence of intervals such that $\sum_n \lambda(I_n) = \infty$ then (I_n) is a BC sequence. Assumptions (a) or (b) of Kim are not necessary for Gouëzel's result. Gouëzel uses renewal theory and obtains BC results but not sBC results.

In the setting of continuous-time systems, Maucourant [25] considers geodesic flows on hyperbolic manifolds of finite volume. He proves a BC result for nested balls in this context.

In this paper, we prove sBC results for intervals satisfying a bounded ratio condition for $1D$ Gibbs-Markov maps. We use this result to establish the dBC property for sequences of nested intervals in the setting of nonuniformly expanding $1D$ systems modeled by Young towers. More precisely, our dBC results are formulated for sequences of nested intervals $I(n)$ with center x_n and length $g(n)$. Our assumption that the intervals are nested implies that $g(n)$ is a decreasing sequence. In specific situations, one often sets $g(n) = n^{-\beta}$ for some $0 \leqslant \beta \leqslant 1$. Many nonuniformly expanding $1D$ maps can be modeled by Young towers. If $(M, \mathcal{B}, \mu, T)$ is a $C^{1+\varepsilon}$ dynamical system on a compact interval M such that μ is ergodic, $\mu \ll \lambda$, and μ has a positive Lyapunov exponent, then the system can be modeled by a Young tower (personal communication by José Alves and Henk Bruin; see also [2, 3, 7]). The results of this paper therefore apply to such maps.

This paper is organized as follows. In Section 2 we describe our main results: Theorems 1, 2, and 3. We prove these results in Sections 3 and 4. Finally, we briefly discuss applications to the extreme value theory of dynamical systems in Section 5.

2. Setting and statements of results

2.1. **Gibbs-Markov maps.** We first describe $1D$ Gibbs-Markov maps and then show that for such maps, sequences of intervals satisfying a bounded ratio criterion have the sBC property. The base map of a Young tower is a Gibbs-Markov system and our result for such systems, Theorem 1, will play a crucial role in the proof of Theorems 2 and 3.

Let $(X, \mathcal{B}, m)$ be a Lebesgue probability space. Let $\mathcal{P}$ be a countable measurable partition of X such that $m(\alpha) > 0$ for all $\alpha \in \mathcal{P}$.

Definition 2.1. A measure-preserving map $T : X \to X$ is said to be a ***Markov map*** if the following are satisfied.

(1) ($\mathcal{P}$ generates $\mathcal{B}$) We have $\sigma(\{T^{-i}(\alpha) : \alpha \in \mathcal{P},\ i \in \mathbb{Z}^+\}) = \mathcal{B} \pmod m$, where $\sigma(\cdot)$ denotes the σ-algebra generated by its argument.
(2) (Markov property) For all $\alpha, \beta \in \mathcal{P}$, if $m(T(\alpha) \cap \beta) > 0$ then $\beta \subset T(\alpha) \pmod m$.
(3) (local invertibility) For all $\alpha \in \mathcal{P}$, $T|\alpha$ is invertible.

For $n \in \mathbb{N}$, let $\mathcal{P}_n$ be the partition of X defined by

$$\mathcal{P}_n = \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) = \left\{ \bigcap_{i=0}^{n-1} T^{-i}(\alpha_i) : \alpha_i \in \mathcal{P} \text{ for } 0 \leqslant i \leqslant n-1 \right\}.$$

Define

$$J_T = \frac{d(m \circ T)}{dm}.$$

Definition 2.2. The quintet $(X, \mathcal{B}, m, T, \mathcal{P})$ is said to be a ***Gibbs-Markov system*** if T is a Markov map and the following properties also hold.

(H1) (full branches) For all $\alpha \in \mathcal{P}$, $T(\alpha) = X \pmod m$.

(H2) (uniform expansion) There exists $K_1 > 0$ and $\gamma_1 \in (0,1)$ such that $m(\alpha) \leqslant K_1 \gamma_1^n$ for all $n \in \mathbb{N}$ and $\alpha \in \mathcal{P}_n$.

(H3) (distortion control) There exists $K_2 > 0$ and $\gamma_2 \in (0,1)$ such that for all $n \in \mathbb{N}$ and $\alpha \in \mathcal{P}_n$, we have

$$\left| \log \left(\frac{J_T(x)}{J_T(y)} \right) \right| \leqslant K_2 \gamma_2^n \tag{2}$$

for all $x, y \in \alpha$.

Remark 2.3. Some authors weaken (H1) in the definition of Gibbs-Markov systems by requiring that $m(T(\alpha)) > K > 0$ for some K independent of α.

Definition 2.4. The Gibbs-Markov system $(X, \mathcal{B}, m, T, \mathcal{P})$ is said to be a **$1D$ Gibbs-Markov system** if X is a compact interval and $\mathcal{P}$ is a partition of X into subintervals.

Remark 2.5. A $1D$ Gibbs-Markov system $(X, \mathcal{B}, m, T, \mathcal{P})$ need not satisfy the condition that $1/T'$ is of bounded variation. For example, let $0 < \varepsilon < 1/2$ and for $n \geqslant 2$ suppose that (I_n) is a sequence of contiguous disjoint open intervals such that $\lambda(I_n) = n^{-1-\varepsilon}$ and $\bigcup_{n \geqslant 2} \bar{I}_n = X$, an interval of length L. Suppose T maps each I_n onto X with bounded distortion so that $T'(x) \approx L n^{1+\varepsilon}$ for each $x \in I_n$. On I_n we allow T' to vary. For each I_n, let $x_1^n, x_2^n, \ldots, x_{m_n}^n$ be a sequence of decreasing points in I_n such that $T'(x_j^n) = L(n^{1+\varepsilon} + (-1)^j n)$. Note that $T'(x_j^n) > 0$ and

$$\left| \frac{1}{T'(x_j^n)} - \frac{1}{T'(x_{j+1}^n)} \right| > 2L^{-1} n^{-1-2\varepsilon}.$$

We set $m_n = n$ so that the variation of $1/T'$ on I_n is greater than $2L^{-1} n^{-2\varepsilon}$. As $\varepsilon < 1/2$, $1/T'$ is not of bounded variation on X. Note that T may be constructed to be piecewise $C^{1+\alpha}$ for any $0 < \alpha < 1$.

Now let X be a compact interval. A map $T : X \to X$ is said to be piecewise-differentiable if there exists a countable partition $\mathcal{P}$ of X into intervals with disjoint interiors such that for all $I \in \mathcal{P}$, T is differentiable on the interior of I. A piecewise-differentiable map $T : X \to X$ is said to be uniformly expanding if there exists $K > 1$ such that $|T'(x)| \geqslant K$ for all x at which $T'(x)$ exists. We similarly define piecewise-C^k maps for $k \geqslant 2$.

For certain piecewise-differentiable uniformly expanding maps, Kim [21, Theorem 2.1] establishes the sBC property for sequences of intervals. His result can be more usefully stated as

Proposition 2.6 ([21])**.** *Suppose T is a piecewise-differentiable uniformly expanding map of the compact interval X and suppose that T admits a unique absolutely continuous invariant probability measure μ with density bounded away from 0. Assume that there exists a summable sequence $(\kappa(n))_{n=1}^{\infty}$ and $C > 0$ such that for all $f \in L^1(\mu)$ and $\psi \in \mathrm{BV}(X)$, we have*

$$\left| \int_X (f \circ T^n)(\psi)\, d\mu - \left(\int_X f\, d\mu \right) \left(\int_X \psi\, d\mu \right) \right| \leqslant C\kappa(n) \|f\|_1 \|\psi\|_{\mathrm{BV}}. \tag{3}$$

If (A_n) is a sequence of intervals in X and $\sum_{n=0}^{\infty} \mu(A_n) = \infty$ then (A_n) is an sBC sequence.

The proof is the same as that of [21, Theorem 2.1]. Note that Gibbs-Markov maps need not satisfy the decay of correlations assumptions of Proposition 2.6 for observations in the Banach spaces $L^1(\mu)$ and $BV(X)$. However, if $T : X \to X$ is C^2 and satisfies certain estimates on its second derivative then Kim's assumptions hold. We state this as a proposition.

Proposition 2.7. *Suppose that $(X, \mathcal{B}, m, T, \mathcal{P})$ is a piecewise-C^2 1D Gibbs-Markov system for which there exists $L > 0$ such that*

$$\sup_{\alpha \in \mathcal{P}} \sup_{x \in \overline{\alpha}} \frac{|T''(x)|}{T'(x)^2} \leqslant L < \infty. \tag{4}$$

Let $(A_n)_{n=0}^{\infty}$ be a sequence of intervals in X. If $\sum_{n=0}^{\infty} m(A_n) = \infty$, then (A_n) is an sBC sequence.

Proof of Proposition 2.7. Condition (4) is sometimes called the Adler property. It enables us to show that $g := 1/T'$ is of bounded variation. Rychlik [29] has shown that for piecewise-differentiable uniformly expanding maps with g of bounded variation, correlations decay exponentially; that is, (3) holds with $\kappa(n)$ decaying exponentially (see also [4, 19]). Kim uses the result of Rychlik to establish the sBC property for sequences of intervals in the setting of piecewise-differentiable uniformly expanding maps with g of bounded variation, although his proof is valid if $(\kappa(n))$ is summable.

To see that g is of bounded variation, let $x, y \in \alpha \in \mathcal{P}$. Using (4) and (H3), we have

$$|g(x) - g(y)| = \left| \frac{T'(y) - T'(x)}{T'(x)T'(y)} \right| \leqslant \int_x^y \frac{|T''(s)|}{|T'(x)T'(y)|}\, ds = \int_x^y \left(\frac{|T''(s)|}{T'(s)^2} \right) \left(\frac{T'(s)^2}{|T'(x)T'(y)|} \right) ds \leqslant KL|x - y|.$$

Using the distortion estimate (H3) again, for every $\alpha \in \mathcal{P}$ and for every $x \in \alpha$, we have

$$e^{-K_2} \left(\frac{\lambda(X)}{\lambda(\alpha)} \right) \leqslant |T'(x)| \leqslant e^{K_2} \left(\frac{\lambda(X)}{\lambda(\alpha)} \right)$$

where λ denotes Lebesgue measure on $\mathbb{R}$. Consequently, if $x \in \alpha \in \mathcal{P}$ and $y \in \beta \in \mathcal{P}$, then

$$|g(x) - g(y)| \leqslant \frac{1}{|T'(x)|} + \frac{1}{|T'(y)|} \leqslant K(\lambda(\alpha) + \lambda(\beta)).$$

These observations show that g is of bounded variation if the Adler condition holds and hence that (A_n) is an sBC sequence. ∎

We now state a result for 1D Gibbs-Markov systems (not necessarily satisfying $1/T'$ in $BV(X)$) with a bounded ratio restriction on the sequence of intervals (A_n). The proof of Theorem 1 is given in Section 3.

Theorem 1. *Let X be a compact interval and let $\mathcal{P}$ be a countable partition of X into subintervals. Suppose that $(X, \mathcal{B}, m, T, \mathcal{P})$ is a 1D Gibbs-Markov system. Let $(A_n)_{n=0}^{\infty}$ be a sequence of intervals in X for which there exists $C_1 > 0$ such that $m(A_j) \leqslant C_1 m(A_i)$ for all $j \geqslant i \geqslant 0$. If $\sum_{n=0}^{\infty} m(A_n) = \infty$, then (A_n) is an sBC sequence.*

2.2. **Young towers for 1-dimensional maps.** Let $M \subset \mathbb{R}$ be a compact interval and let μ be a probability measure on M with σ-algebra $\mathcal{B}$. Suppose that $\mu \ll \lambda$. Suppose that $T : M \to M$ preserves μ and that $(M, \mathcal{B}, \mu, T)$ is ergodic. We describe what it means for $(M, \mathcal{B}, \mu, T)$ to be modeled by a Young tower. See [31, 32] for more details.

There exists a subinterval $\Lambda_0 \subset M$, a countable partition $\{\Lambda_i : i \in \mathbb{N}\}$ of Λ_0 into subintervals, and a roof function $R : \Lambda_0 \to \mathbb{N}$ such that $R \in L^1(\lambda)$, $R|\Lambda_i \equiv L_i$ where $L_i \in \mathbb{N}$, and $T^{L_i}(\Lambda_i) = \Lambda_0$. The Young tower Δ is defined by

$$\Delta = \bigcup_{i=1}^{\infty} \bigcup_{l=0}^{L_i - 1} \Lambda_{i,l}$$

where $\Lambda_{i,l} := (\Lambda_i, l)$. The tower map $F : \Delta \to \Delta$ is given by

$$F(x,l) = \begin{cases} (x, l+1), & \text{if } x \in \Lambda_i \text{ and } l < L_i - 1; \\ (T^{L_i}(x), 0), & \text{if } x \in \Lambda_i \text{ and } l = L_i - 1. \end{cases}$$

We call Λ_0 the base of the tower and we call R the return time function associated with Λ_0. The map $f := F^R : \Lambda_0 \to \Lambda_0$ is called the base map.

The following properties of the base map follow from the properties of a Young tower. The map f is a $1D$ Gibbs-Markov map with respect to the partition $\{\Lambda_i\}$ and admits an ergodic invariant probability measure m. The measure m is equivalent to λ in the sense that the Radon-Nikodym derivative $dm/d\lambda$ satisfies

$$C^{-1} \leqslant \frac{dm}{d\lambda} \leqslant C \tag{5}$$

for some constant $C > 0$. The measure m lifts to an invariant measure ν for F on Δ by defining

$$\nu(E) = \frac{m(F^{-l}(E))}{\int_{\Lambda_0} R \, dm}$$

for $E \subset \Lambda_{i,l}$ and extending the definition in the natural way to subsets of Δ.

The system (Δ, F) is an extension of (M, T) in the sense that the projection $\pi : \Delta \to M$ defined by $\pi(x,l) = T^l(x)$ satisfies $\pi \circ F = T \circ \pi$. The invariant measure μ for T is given by $\mu = \pi_* \nu$. We summarize the setting described above by saying that $(M, \mathcal{B}, \mu, T)$ is a $1D$ system modeled by a Young tower Δ.

Theorem 2. *Let $(M, \mathcal{B}, \mu, T)$ be a $1D$ dynamical system. Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a decreasing function and $(I(n))_{n=0}^{\infty}$ a nested sequence of closed intervals in M such that $I(n)$ has length $g(n)$ for all $n \in \mathbb{Z}^+$ and $\sum_{n=0}^{\infty} g(n) = \infty$. If $(M, \mathcal{B}, \mu, T)$ is modeled by a Young tower Δ such that*

$$J_\infty := \bigcap_{n=0}^{\infty} I(n) \subset \pi(\Delta),$$

then the following hold.

(1) If J_∞ is an interval, then $(I(n))$ is an sBC sequence.

(2) There exists a set $\Gamma \subset M$, $\mu(\Gamma) = 1$, such that if $p = J_\infty \in \Gamma$ then $(I(n))$ is a dBC sequence with respect to μ satisfying

$$\varliminf_{n\to\infty} \frac{S_n(x)}{E_n} \geqslant 1 \quad a.s.,$$

where $E_n = \sum_{j=0}^{n-1} \mu(I(j))$.

Remark 2.8. If J_∞ is an interval then the conclusion of Theorem 2 follows immediately from the Birkhoff ergodic theorem.

Remark 2.9. It is easy to show that for any $p \in M$ and for every k and l, either $\Lambda_{k,l} \cap \pi^{-1}(p) = \emptyset$ or $\Lambda_{k,l} \cap \pi^{-1}(p)$ consists of a single point $\hat{p}_{k,l}$. The set Γ consists of those points $p \in M$ such that $\hat{p}_{k,l} \in \operatorname{int}(\Lambda_{k,l})$ for all k and l for which $\hat{p}_{k,l}$ is defined and

$$\lim_{r\to 0} \frac{\mu(B(p,r))}{2r} = \frac{d\mu}{d\lambda}(p) > 0,$$

where $B(p,r)$ is the ball (interval) centered at p of radius r.

The following dBC result for nested intervals centered at a point $p \in M$ does not require that $p \in \Gamma$.

Theorem 3. *Suppose that $(M, \mathcal{B}, \mu, T)$ is a $1D$ system. Let $p \in M$ and let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a decreasing function such that $\sum_{n=0}^{\infty} g(n) = \infty$. For $n \in \mathbb{Z}^+$, let $I(n)$ denote the closed interval centered at p of length $g(n)$. If $(M, \mathcal{B}, \mu, T)$ may be modeled by a Young tower Δ such that $p \in \pi(\Delta)$, then $(I(n))$ is a dBC sequence with respect to Lebesgue measure λ.*

As mentioned earlier, if $(M, \mathcal{B}, \mu, T)$ is a $C^{1+\varepsilon}$ dynamical system on a compact interval M such that μ is ergodic, $\mu \ll \lambda$, and μ has a positive Lyapunov exponent, then the system may be modeled by a Young tower (personal communication by José Alves and Henk Bruin; see also [2, 3, 7]). In fact, in certain systems (see Section 2.3), for all but finitely many points p there exists a tower Δ (which may depend on p) such that $p \in \pi(\Delta)$. This implies the following corollary.

Corollary 4. *Suppose that $(M, \mathcal{B}, \mu, T)$ is a $C^{1+\varepsilon}$ dynamical system on a compact interval M such that the invariant probability measure μ is ergodic and absolutely continuous with respect to Lebesgue measure. If $(M, \mathcal{B}, \mu, T)$ has a positive Lyapunov exponent, then Theorems 2 and 3 apply to $(M, \mathcal{B}, \mu, T)$.*

2.3. **1-dimensional maps.** Theorem 2 applies to many classes of 1-dimensional maps. Here we give a partial list.

(1) Pomeau-Manneville intermittent-type maps (such as Liverani-Saussol-Vaienti maps) [24, 27]. See [15, 21] for related results. In this setting the only points p that cannot be included in $\pi(\Delta)$ for some Young tower Δ are the indifferent fixed points. For Liverani-Saussol-Vaienti maps (1), this is the origin $p = 0$.
(2) Certain classes of unimodal maps. In the setting of logistic maps $T_a(x) = 1 - ax^2$, $T : [-1, 1] \to [-1, 1]$ with a a Benedicks-Carleson parameter [5], the only points p which may not be included in $\pi(\Delta)$ for some Young tower Δ are the points $p = 0$ and $p = T(0)$.
(3) Certain classes of multimodal maps; see Bruin *et. al.* [6]. In the setting of [6] there are only finitely points p that may not be included in $\pi(\Delta)$ for some Young tower Δ (personal communication by Henk Bruin). Such points are critical points and perhaps a few iterates of critical points.
(4) A class of nonuniformly expanding circle maps; see Young [32, Section 6]. Let $T : \mathbb{S}^1 \to \mathbb{S}^1$ be a map of degree $d > 1$ such that T is C^1 on $\mathbb{S}^1$ and C^2 on $\mathbb{S}^1 \setminus \{0\}$, $T' > 1$ on $\mathbb{S}^1 \setminus \{0\}$, $T(0) = 0$, $T'(0) = 1$, and for some $0 < \alpha < 1$, we have $-xT''(x) \sim |x|^\alpha$ for $x \neq 0$. In this setting the only point which may not be included in $\pi(\Delta)$ for some Young tower Δ is $p = 0$.

3. Proof of Theorem 1

We prove Theorem 1 by using a sufficient condition, Proposition 3.1, for the sBC property given in [22, 30]. This sufficient condition has also been used by Chernov and Kleinbock [8] and by Kim [21]. From the conclusion of Proposition 3.1 it is easy to see that, upon dividing both sides of equation (7) by $\sum_{i=0}^{N-1} m(B_i)$ and taking the limit as $N \to \infty$ we obtain the sBC property for the sequence (B_n).

Proposition 3.1. *Let $(X, \mathcal{B}, m)$ be a probability space and let $(B_n)_{n=0}^{\infty}$ be a sequence of measurable subsets of X such that $\sum_{n=0}^{\infty} m(B_n) = \infty$. If there exists a constant $C_2 > 0$ such that for all $N \geqslant M \geqslant 0$*

$$\sum_{i,j=M}^{N} (m(B_i \cap B_j) - m(B_i)m(B_j)) \leqslant C_2 \sum_{i=M}^{N} m(B_i), \tag{6}$$

then for every $\varepsilon > 0$, we have

$$\sum_{i=0}^{N-1} \mathbf{1}_{B_i}(x) = \sum_{i=0}^{N-1} m(B_i) + \mathcal{O}\left(\left(\sum_{i=0}^{N-1} m(B_i)\right)^{\frac{1}{2}} \log^{\frac{3}{2}+\varepsilon}\left(\sum_{i=0}^{N-1} m(B_i)\right)\right) \tag{7}$$

for m-a.e. $x \in X$.

Remark 3.2. The implied constant in the error estimate

$$\mathcal{O}\left(\left(\sum_{i=0}^{N-1} m(B_i)\right)^{\frac{1}{2}} \log^{\frac{3}{2}+\varepsilon}\left(\sum_{i=0}^{N-1} m(B_i)\right)\right)$$

is not uniform; it is a function of x.

Our proof uses the fact that Gibbs-Markov maps are exponentially continued fraction mixing [1, page 164] in the sense that there exists $\tau \in (0,1)$ and a constant $K_3 > 0$ such that

$$|m(\alpha \cap T^{-(n+k)}(\beta) - m(\alpha)m(\beta)| \leqslant K_3\tau^n m(\alpha)m(\beta)$$

for all measurable $\beta \in \mathcal{B}$, $\alpha \in \mathcal{P}_k$.

Proof of Theorem 1. Throughout this proof C will be used to denote a generic constant, the value of which may change from line to line. Set $B_i = T^{-i}(A_i)$ in (6). Notice that if $j \geqslant i$, then $T^{-i}(A_i) \cap T^{-j}(A_j) = T^{-i}(A_i \cap T^{-(j-i)}(A_j))$. Since T preserves m, we have $m(T^{-i}(A_i \cap T^{-(j-i)}(A_j))) = m(A_i \cap T^{-(j-i)}(A_j))$. We therefore estimate

$$\begin{aligned}\sum_{i,j=M}^{N} m(B_i \cap B_j) - m(B_i)m(B_j) = 2\sum_{i=M}^{N}\sum_{j=i+1}^{N} m(A_i \cap T^{-(j-i)}(A_j)) - m(A_i)m(A_j) \\ + \sum_{i=M}^{N} m(A_i) - (m(A_i))^2.\end{aligned} \tag{8}$$

For the diagonal terms, we have the straightforward estimate

$$\sum_{i=M}^{N}\left(m(A_i) - (m(A_i))^2\right) \leqslant \sum_{i=M}^{N} m(A_i). \tag{9}$$

Now assume that $j > i$. We estimate $m(A_i \cap T^{-(j-i)}(A_j)) - m(A_i)m(A_j)$.

Let $\mathcal{V}_{i,j} = \{\alpha \in \mathcal{P}_{\lceil (j-i)/2\rceil} : \alpha \cap A_i \neq \emptyset\}$. Let $N(i)$ be the largest integer such that A_i intersects at most 2 partition elements of $\mathcal{P}_{\lceil (j-i)/2\rceil}$ for $j-i < N(i)$. If $N(i) > j-i > 1$ and $j-i$ is even, we have the estimate

$$\begin{aligned} m(A_i \cap T^{-(j-i)}A_j) - m(A_i)m(A_j) \leqslant m(A_i \cap T^{-(j-i)}A_j) &= m(A_i \cap T^{-(j-i)/2}(T^{-(j-i)/2}A_j)) \\ &\leqslant 2C\gamma_1^{(j-i)/2} m(T^{-(j-i)/2}A_j) \\ &= 2C\gamma_1^{(j-i)/2} m(A_j) \\ &\leqslant C\gamma_1^{(j-i)/2} m(A_i). \end{aligned}$$

This holds because

(1) using (H1)–(H3), for each $\alpha \in \mathcal{V}_{i,j}$ we have

$$m(\alpha \cap T^{-(j-i)/2}(T^{-(j-i)/2}A_j)) \leqslant Cm(\alpha)m(T^{-(j-i)/2}A_j)$$

and $m(\alpha) \leqslant K_1\gamma_1^{(j-i)/2}$,

(2) $\#\mathcal{V}_{i,j} \leqslant 2$ since $j-i < N(i)$, and

(3) $m(A_j) \leqslant Cm(A_i)$ by assumption.

If $j-i$ is odd we estimate

$$m(A_i \cap T^{-(j-i)}A_j) - m(A_i)m(A_j) \leq m(A_i \cap T^{-(j-i)}A_j) = m(A_i \cap T^{-(j-i+1)/2}(T^{-(j-i-1)/2}A_j))$$

and proceed as before. In particular, if $j-i=1$ we use the simple estimate $m(A_i \cap T^{-(j-i)}A_j) \leqslant m(A_i)$. Thus

$$\sum_{j=i+1}^{i+N(i)-1} m(A_i \cap T^{-(j-i)}A_j) - m(A_i)m(A_j) \leq Cm(A_i).$$

We now consider $j-i \geq N(i)$. We can no longer assume that A_i intersects at most 2 elements of $\mathcal{P}_{\lceil (j-i)/2\rceil}$. In this case the collection $\mathcal{V}_{i,j}$ induces a partition of A_i. Define

$$\mathcal{V}^1_{i,j} = \{\alpha \in \mathcal{V}_{i,j} : \alpha \subset (\inf(A_i), \sup(A_i))\}.$$

Since $\mathcal{P}$ consists of subintervals, $\mathcal{V}^2_{i,j} := \mathcal{V}_{i,j} \setminus \mathcal{V}^1_{i,j}$ contains at most 2 elements. For $k = 1,2$ define

$$Q^k_{i,j} = \bigcup\{\alpha : \alpha \in \mathcal{V}^k_{i,j}\}.$$

We have

$$\sum_{j=i+N(i)}^{N} m(A_i \cap T^{-(j-i)}A_j) - m(A_i)m(A_j)$$

$$= \sum_{j=i+N(i)}^{N} m(A_i \cap Q^1_{i,j} \cap T^{-(j-i)}A_j) - m(A_i)m(A_j)$$

$$+ \sum_{j=i+N(i)}^{N} m(A_i \cap Q^2_{i,j} \cap T^{-(j-i)}A_j) - m(A_i)m(A_j)$$

$$\leqslant \sum_{j=i+N(i)}^{N} m(A_i \cap Q^1_{i,j} \cap T^{-(j-i)}A_j) - m(A_i \cap Q^1_{i,j})m(A_j) \tag{10a}$$

$$+ \sum_{j=i+N(i)}^{N} m(A_i \cap Q^2_{i,j} \cap T^{-(j-i)}A_j) - m(A_i \cap Q^2_{i,j})m(A_j). \tag{10b}$$

We estimate (10a) first. For this we will use the exponential continued fraction mixing estimate of Aaronson [1, page 164]. There exists $C > 0$ and $\tau \in (0,1)$, both independent of i and j, such that for $\alpha \in \mathcal{V}^1_{i,j}$, we have

$$|m(\alpha \cap T^{-(j-i)}(A_j)) - m(\alpha)m(A_j)| \leqslant C\tau^{(j-i)/2} m(\alpha)m(A_j).$$

Thus

$$\sum_{j=i+N(i)}^{N} \sum_{\alpha \in \mathcal{V}^1_{i,j}} m(\alpha \cap T^{-(j-i)}A_j) - m(\alpha)m(A_j) \leqslant \sum_{j=i+N(i)}^{N} \sum_{\alpha \in \mathcal{V}^1_{i,j}} C\tau^{(j-i)/2} m(\alpha)m(A_j)$$

$$\leqslant \sum_{j=i+N(i)}^{N} C\tau^{(j-i)/2} m(A_i)m(A_j).$$

For (10b) we note that if $\beta \in \mathcal{V}^2_{i,j}$ then $m(\beta) \leqslant K_1 \gamma_1^{\lceil (j-i)/2 \rceil}$. Consider the partition of β induced by $\mathcal{P}_{j-i}$.

For $\omega \in \mathcal{P}_{j-i}$ such that $\omega \subset \beta$, $T^{j-i}(\omega) = X$ by (H1), so the distortion estimate (H3) and uniform expansion (H2) give

$$m(\beta \cap T^{-(j-i)}(A_j)) = \sum_{\substack{\omega \in \mathcal{P}_{j-i} \\ \omega \subset \beta}} m(\omega \cap T^{-(j-i)}(A_j)) \leqslant C\gamma_1^{(j-i)/2} m(A_j)$$

and hence

$$\sum_{j=i+N(i)}^{N} m(A_i \cap Q^2_{i,j} \cap T^{-(j-i)}A_j) - m(A_i \cap Q^2_{i,j})m(A_j) \leqslant \sum_{j=i+N(i)}^{N} 2C\gamma_1^{(j-i)/2} m(A_j).$$

This conclude the proof as $m(A_j) \leqslant Cm(A_i)$. ■

4. Proof of Theorems 2 and 3

4.1. **Relating base dynamics to tower dynamics and preliminaries.** For $x \in \Lambda_0$ and $n \in \mathbb{N}$, we define

$$R_n(x) = \sum_{i=0}^{n-1} R(f^i(x)).$$

As a consequence of the Birkhoff ergodic theorem, we have

Lemma 4.1. *Assume the setting of Theorem 2 and define*

$$\langle R \rangle = \int_{\Lambda_0} R \, dm.$$

Then for m a.e. $x \in \Lambda_0$, we have

$$\lim_{n\to\infty} \frac{R_n(x)}{n} = \langle R\rangle. \tag{11}$$

The following elementary lemmas will prove to be useful.

Lemma 4.2. *Suppose $g : \mathbb{R}^+ \to \mathbb{R}^+$ is decreasing and $\sum_{i=0}^{\infty} g(i) = \infty$.*

(A) *For all $a > 0$ we have*

$$\frac{\int_0^{(1+a)n} g(t)\,dt}{\int_0^n g(t)\,dt} \leqslant 1 + a$$

for all $n \in \mathbb{N}$.

(B)

$$\lim_{n\to\infty} \frac{\int_0^n g(t)\,dt}{\sum_{i=0}^{n-1} g(i)} = 1.$$

Proof of Lemma 4.2. For part (A), observe that since g is decreasing,

$$\int_0^{(1+a)n} g(t)\,dt = \int_0^n g(t)\,dt + \int_n^{(1+a)n} g(t)\,dt \leqslant \int_0^n g(t)\,dt + (an)g(n) \leqslant (1+a)\int_0^n g(t)\,dt.$$

For part (B), the bound

$$\sum_{i=0}^{n-1} g(i) \geqslant \int_0^n g(t)\,dt \geqslant \sum_{i=1}^{n} g(i)$$

implies the result since $\sum_{i=0}^{\infty} g(i) = \infty$. ∎

4.2. **Proof of Theorem 2.** Assume that

$$\bigcap_{n=0}^{\infty} I(n) = \{p\}.$$

The case that $\bigcap_{n=0}^{\infty} I(n)$ is an interval follows from the Birkhoff ergodic theorem. We will consider partition elements $\Lambda_{k,l}$ of the tower such that $\pi^{-1}(p)\cap\Lambda_{k,l} \neq \emptyset$. Since $\pi^*\nu = \mu$, these are the only partition elements we need to consider to determine whether $F^n(x,0) \in \pi^{-1}(I(n))$ infinitely often (and hence whether $T^n(x) \in I(n)$ infinitely often). For all k and l such that $\pi^{-1}(p) \cap \Lambda_{k,l} \neq \emptyset$, let $\hat{p}_{k,l}$ denote the point of intersection. We assume that $\hat{p}_{k,l} \in \operatorname{int}(\Lambda_{k,l})$ for all k and l for which $\hat{p}_{k,l}$ exists. The notation $\lfloor r\rfloor$ will denote the integer part of a real number.

We first consider the sequence $\Lambda_{k,l} \cap \pi^{-1}(I(n))$ for a fixed partition element $\Lambda_{k,l}$. For $n \in \mathbb{Z}^+$ define

$$A'_n = \Lambda_{k,l} \cap \pi^{-1} I(\lfloor n\langle R\rangle\rfloor), \qquad G'_n = F^{-l}(A'_n)$$

and for $n \in \mathbb{N}$ let

$$\alpha(n) = \sum_{j=0}^{n-1} m(G'_j).$$

For $x \in \Lambda_0$ and $n \in \mathbb{N}$, define

$$\hat{S}(n,x) = \#\{j < n : F^j(x,0) \in \Lambda_{k,l} \cap \pi^{-1}(I(j))\}.$$

Step 1. We relate the recurrence properties of F to those of $f := T^R$. We claim that

$$\lim_{n\to\infty} \frac{\hat{S}(n,x)}{\alpha(\lfloor n\langle R\rangle^{-1}\rfloor)} = 1$$

for m a.e. $x \in \Lambda_0$. As the proof of the claim proceeds, finitely many restrictions will be placed on x, each of which is satisfied by a set of full measure. First, assume that x satisfies (11). Ergodicity relates the clock associated with the tower map F to the clock associated with the return map f. For $\varepsilon \in \mathbb{R}$ with small modulus, define the sets

$$A_{n,\varepsilon} = \Lambda_{k,l} \cap \pi^{-1}(I(\lfloor n(\langle R\rangle + \varepsilon)\rfloor)), \qquad G_{n,\varepsilon} = F^{-l}(A_{n,\varepsilon})$$

and the sums

$$\mathcal{S}_\varepsilon(n,x) = \sum_{j=0}^{n-1} \mathbf{1}_{G_{j,\varepsilon}} \circ f^j(x), \qquad \mathcal{E}_\varepsilon(n) = \sum_{j=0}^{n-1} m(G_{j,\varepsilon}).$$

For sequences (u_n) and (v_n) in $\mathbb{R}$, we write $u_n \approx v_n$ if

$$\lim_{n\to\infty} \frac{u_n}{v_n} = 1.$$

Since $\hat{p}_{k,l} \in \operatorname{int}(\Lambda_{k,l})$, we have $\mathcal{E}_\varepsilon(n) \to \infty$ as $n \to \infty$. Theorem 1 gives the sBC property for the base transformation: for m a.e. $x \in \Lambda_0$, we have

$$\lim_{n\to\infty} \frac{\mathcal{S}_\varepsilon(n,x)}{\mathcal{E}_\varepsilon(n)} = 1. \tag{12}$$

We now examine $\hat{S}(n,x)$. Define $q(n,x)$ by

$$R_{q(n,x)}(x) \leqslant n < R_{q(n,x)+1}(x).$$

Observe that $\hat{S}(n,x) - \hat{S}(R_{q(n,x)}, x) \in \{0,1\}$. This is so because the levels of the Young tower are pairwise disjoint and so the orbit of $(x,0)$ must enter $\Lambda_{k,l}$ in order to increment $\hat{S}$ and this can happen at most once from time $R_{q(n,x)}$ to time n. Thus it suffices to examine $\hat{S}(R_{q(n,x)}, x)$. Using (11), for $\varepsilon > 0$ small we obtain

$$\left(\frac{\mathcal{S}_\varepsilon(q(n,x),x) + \psi(x,\varepsilon)}{\mathcal{E}_\varepsilon(q(n,x))}\right)\left(\frac{\mathcal{E}_\varepsilon(q(n,x))}{\alpha(q(n,x))}\right) \leqslant \frac{\hat{S}(R_{q(n,x)},x)}{\alpha(q(n,x))} \leqslant \left(\frac{\mathcal{S}_{-\varepsilon}(q(n,x),x) + \zeta(x,\varepsilon)}{\mathcal{E}_{-\varepsilon}(q(n,x))}\right)\left(\frac{\mathcal{E}_{-\varepsilon}(q(n,x))}{\alpha(q(n,x))}\right).$$

where ψ and ζ are independent of n. Using (12) and Lemma 4.2, this implies

$$\lim_{n\to\infty} \frac{\hat{S}(n,x)}{\alpha(q(n,x))} = 1$$

and therefore another application of Lemma 4.2 gives

$$\lim_{n\to\infty} \frac{\hat{S}(n,x)}{\alpha(\lfloor n\langle R\rangle^{-1}\rfloor)} = 1$$

since $q(n,x) \approx n\langle R\rangle^{-1}$.

Step 2. We claim that

$$\alpha(\lfloor n\langle R\rangle^{-1}\rfloor) \approx \sum_{j=0}^{n-1} \nu(\hat{I}_{k,l}(j))$$

where

$$\hat{I}_{k,l}(j) = \Lambda_{k,l} \cap \pi^{-1}(I(j)).$$

This follows from a change of variable argument. Let $\tilde{p}_{k,l} \in \Lambda_k$ satisfy $T^l(\tilde{p}_{k,l}) = p$. Defining $\rho = dm/d\lambda$, we have

$$\begin{aligned}
\alpha(\lfloor n\langle R\rangle^{-1}\rfloor) = \sum_{j=0}^{\lfloor n\langle R\rangle^{-1}\rfloor - 1} m(G'_j) &\approx \frac{\rho(p)}{|DT^l(\tilde{p}_{k,l})|} \sum_{j=0}^{\lfloor n\langle R\rangle^{-1}\rfloor - 1} g(j\langle R\rangle) \\
&\approx \frac{\rho(p)}{|DT^l(\tilde{p}_{k,l})|} \int_0^{n\langle R\rangle^{-1}} g(t\langle R\rangle)\,dt \\
&= \frac{\rho(p)}{\langle R\rangle|DT^l(\tilde{p}_{k,l})|} \int_0^{n} g(u)\,du \\
&\approx \frac{\rho(p)}{\langle R\rangle|DT^l(\tilde{p}_{k,l})|} \sum_{j=0}^{n-1} g(j) \\
&\approx \sum_{j=0}^{n-1} \nu(\hat{I}_{k,l}(j)).
\end{aligned}$$

Steps (1) and (2) imply that

$$\hat{S}(n,x) \approx \sum_{j=0}^{n-1} \nu(\hat{I}_{k,l}(j)).$$

Step 3. We now study the sequence of preimages $(\pi^{-1}(I(n))_{n=0}^{\infty}$ on the whole tower Δ. Since $\pi_*\nu = \mu$, we have

$$\frac{d\mu}{d\lambda}(p) = \sum_{\hat{p}_{k,l} \in \pi^{-1}(p)} \frac{\xi_{k,l}(\hat{p}_{k,l})}{|DT^l(\tilde{p}_{k,l})|},$$

where $\xi_{k,l}$ is the density of the restriction of ν to $\Lambda_{k,l}$ and $\tilde{p}_{k,l} \in \Lambda_k$ satisfies $T^l(\tilde{p}_{k,l}) = p$. Consequently, for every $\delta > 0$, there exists $N(\delta)$ such that the truncated tower $\Delta_{N(\delta)} := \{\Lambda_{k,l} : k \leqslant N(\delta) \text{ and } l < \min\{N(\delta), L_k\}\}$ satisfies

$$\mu(I(n) \cap \pi(\Delta_{N(\delta)})) \geqslant (1-\delta)\mu(I(n))$$

for all n sufficiently large. Now fix $\delta > 0$. Repeat steps (1) and (2) for every $\hat{p}_{k,l} \in \pi^{-1}(p)$. For m a.e. $x \in \Lambda_0$, we have

$$\#\{0 \leqslant j < n : F^j(x,0) \in \hat{I}_{k,l}(j)\} \approx \sum_{j=0}^{n-1} \nu(\hat{I}_{k,l}(j)) \tag{13}$$

for every k and l such that $\Delta_{N(\delta)} \cap \pi^{-1}(p) \neq \emptyset$. For $x \in \Lambda_0$, define

$$U(n,x) = \#\left\{0 \leqslant j < n : F^j(x,0) \in \bigcup_{\hat{I}_{k,l}(j) \subset \Delta_{N(\delta)}} \hat{I}_{k,l}(j)\right\}.$$

Estimate (13) and the fact that the levels of the tower are pairwise disjoint imply the existence of $\kappa(\delta)$ satisfying $1-\delta \leqslant \kappa(\delta) \leqslant 1$ for which m a.e. $x \in \Lambda_0$ satisfies

$$U(n,x) \approx \kappa(\delta) \sum_{j=0}^{n-1} \mu(I(j))$$

for every $\delta > 0$. Define

$$V(n,x) = \#\left\{0 \leqslant j < n : F^j(x,0) \in \bigcup_{\hat{I}_{k,l}(j) \in \Delta} \hat{I}_{k,l}(j)\right\}.$$

Since $\delta > 0$ is arbitrary, we conclude that for m a.e. $x \in \Lambda_0$, we have

$$\varliminf_{n\to\infty} \frac{V(n,x)}{\sum_{j=0}^{n-1} \mu(I(j))} \geqslant 1$$

and therefore

$$\varliminf_{n\to\infty} \frac{S_n(x)}{\sum_{j=0}^{n-1} \mu(I(j))} \geqslant 1 \tag{14}$$

where

$$S_n(x) = \#\{0 \leqslant j < n : T^j(x) \in I(j)\}.$$

The arguments in steps (1)–(3) extend to ν a.e. element of every level of the tower Δ. Relating the dynamics on the tower to those on M, since $\pi \circ F = T \circ \pi$ and $\pi_*\nu = \mu$, we conclude that (14) holds for μ a.e. $x \in M$. This concludes the proof of Theorem 2.

4.3. **Proof of Theorem 3.** Since $p \in \pi(\Delta)$ we may consider a single preimage $\hat{p}_{k,l} \in \pi^{-1}(p)$ and note that $J(n) = \pi^{-1}(I(n)) \cap \Lambda_{k,l}$ has the property that $\sum_n \nu(J_n) = \infty$. We then use the same argument as the one given in steps (1) and (2) of the proof of Theorem 2 to obtain the dBC property. Since we do not have the stronger requirement that $p \in \Gamma$ we cannot obtain the stronger estimate that

$$\varliminf_{n\to\infty} \frac{S_n(x)}{\sum_{j=0}^{n-1} \mu(I(j))} \geqslant 1,$$

merely that

$$\varliminf_{n\to\infty} \frac{S_n(x)}{\sum_{j=0}^{n-1} \lambda(I(j))} \geqslant C > 0.$$

5. Almost-sure results in extreme value theory for dynamical systems

There has been much recent work on the extreme value theory of deterministic dynamical systems [9, 10, 12, 26, 13, 16, 17, 18, 20]. Extreme value theory is concerned with the estimation of extremal events and is a standard statistical methodology used to determine risk in areas such as insurance and weather prediction. There are several texts which describe extreme value theory from a statistical viewpoint; see e.g. [14, 23, 28]. Often the models adopted by statisticians assume that observations are independent identically-distributed random variables, an assumption that is usually not valid for time series of observations of deterministic dynamical systems. Such observations tend to be highly correlated.

If $\phi : X \to \mathbb{R}$ is an integrable observable on a dynamical system (X, μ, T), we define $\phi_j(x) = \phi \circ T^j(x)$ and in turn define (M_n), the sequence of successive maxima, by

$$M_n(x) = \max_{0 \leqslant j \leqslant n-1} \phi_j(x).$$

Most recent research on the extreme value theory of nonuniformly hyperbolic dynamical systems [9, 26, 13, 17] has concentrated on the study of distributional limits of the sequence (M_n). Here the goal is to find scaling constants $a_n > 0$ and $b_n \in \mathbb{R}$ and a nondegenerate distribution $G(x)$ such that

$$\lim_{n\to\infty} \mu(a_n(M_n - b_n) \leqslant x) = G(x).$$

For example, if (X, μ, T) is a nonuniformly expanding map of the interval with exponential decay of correlations as in Collet [9] and $\phi(x) = -\log(d(x, x_0))$, then it is known [13, 20] that for μ a.e. $x_0 \in X$, we have

$$\lim_{n\to\infty} \mu(M_n \leqslant v + \log(n)) = e^{-v}$$

for all $v \in \mathbb{R}$.

Our Borel-Cantelli results allow a description of the almost-sure behavior of the sequence of successive maxima $M_n(x)$, rather than just a distributional description. This allows an estimation of almost-sure upper bounds. In a similar way the law of the iterated logarithm gives an almost-sure upper bound for the rate of growth of scaled Birkhoff sums

$$b_n(x) = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \phi(T^j(x)),$$

namely

$$\varlimsup_{n\to\infty} \frac{b_n(x)}{(n \log(\log(n)))^{1/2}} < C$$

almost surely, in contrast to the central limit theorem, which is a distributional result.

Extreme value theory is related to entrance times to nested balls by the observation that if $\phi(x) = -\log(d(x, x_0))$ and $v \in \mathbb{R}$, $M_n(x) < v + \log(n)$ if and only if $d(T^i(x), x_0) > e^{-v}/n$ for $0 \leqslant i \leqslant n-1$. In the context of Theorem 2, for μ a.e. center x_0 and for every $v \in \mathbb{R}$, we have $\mu(M_n > v + \log(n) \text{ i.o.}) = 1$ since

$$E_n = \sum_{i=0}^{n-1} \mu\left(B\left(x_0, \frac{e^{-v}}{i+1}\right)\right)$$

diverges. By contrast, if $\delta > 0$ then

$$\mu(M_n(x) > v + \log(n) + (1+\delta)\log(\log(n)) \text{ i.o.}) = 0$$

by the classical Borel-Cantelli lemma since

$$\sum_{i=2}^{n-1} \mu\left(B\left(x_0, \frac{e^{-v}}{i(\log(i))^{1+\delta}}\right)\right)$$

converges. Thus the sequence $u_n = v + \log(n) + (1+\delta)\log(\log(n))$ is an almost-sure upper bound for M_n for any $\delta > 0$ and any $v \in \mathbb{R}$. The consideration of the function $\phi(x) = -\log(d(x, x_0))$ is not restrictive; other functions with unique maxima can be considered in this framework and almost-sure upper bounds u_n for the sequence (M_n) can be derived, though the sequence (u_n) will depend upon the form of the function near the maximal point.

Almost-sure lower bounds may be derived as well for maps satisfying (7) of Proposition 3.1. Let $v \in \mathbb{R}$ and define

$$E_n := \sum_{j=0}^{n-1} \mu\left(B\left(x_0, \frac{e^{-v}}{j+1}\right)\right).$$

Suppose that n_l is the largest $j < n$ such that $d(T^j(x), x_0) < e^{-v}/(j+1)$. Thus $S_n(x) = S_{n_l}(x)$, so

$$S_n(x) \leqslant E_{n_l} + C(x)(E_{n_l})^{\frac{1}{2}} \log^{\frac{3}{2}+\varepsilon}(E_{n_l}),$$

where $C(x) > 0$ is the implied constant in (7). Since

$$E_n \leqslant S_n(x) + C(x)E_n^{\frac{1}{2}} \log^{\frac{3}{2}+\varepsilon}(E_n),$$

a simple calculation shows that we have

$$E_n - E_{n_l} \leqslant 2C(x)E_n^{\frac{1}{2}} \log^{\frac{3}{2}+\varepsilon}(E_n).$$

Let $g := d\mu/d\lambda$. Since $M_n(x) > v + \log(n_l + 1)$ and $E_{n_l+1} \approx g(x_0)e^{-v}\log(n_l+1)$, we have

$$M_n(x) > v + \zeta(n_l)g(x_0)^{-1}e^v E_{n_l+1},$$

where $\zeta(n_l)$ is defined by $\log(n_l+1) = \zeta(n_l)g(x_0)^{-1}e^v E_{n_l+1}$ and satisfies $\zeta(n_l) \to 1$ as $n_l \to \infty$. We conclude that given any $\delta > 0$ and any $0 < \varepsilon_1 < 1$, for n sufficiently large we have

$$v + (1-\varepsilon_1)e^v g(x_0)^{-1}\left(E_n - 2C(x)E_n^{\frac{1}{2}} \log^{\frac{3}{2}+\varepsilon}(E_n)\right) < M_n(x) \leqslant v + \log(n) + (1+\delta)\log(\log(n)).$$

References

1. Jon Aaronson, *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, Providence, RI, 1997. MR MR1450400 (99d:28025)
2. José F. Alves, Stefano Luzzatto, and Vilton Pinheiro, *Lyapunov exponents and rates of mixing for one-dimensional maps*, Ergodic Theory Dynam. Systems **24** (2004), no. 3, 637–657. MR MR2060991 (2005h:37083)
3. ———, *Markov structures and decay of correlations for non-uniformly expanding dynamical systems*, Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005), no. 6, 817–839. MR MR2172861 (2006e:37047)
4. Viviane Baladi, *Positive transfer operators and decay of correlations*, Advanced Series in Nonlinear Dynamics, vol. 16, World Scientific Publishing Co. Inc., River Edge, NJ, 2000. MR MR1793194 (2001k:37035)
5. Michael Benedicks and Lennart Carleson, *On iterations of* $1-ax^2$ *on* $(-1,1)$, Ann. of Math. (2) **122** (1985), no. 1, 1–25. MR MR799250 (87c:58058)
6. Henk Bruin, Stefano Luzzatto, and Sebastian Van Strien, *Decay of correlations in one-dimensional dynamics*, Ann. Sci. École Norm. Sup. (4) **36** (2003), no. 4, 621–646. MR MR2013929 (2005a:37061)
7. Henk Bruin and Mike Todd, *Return time statistics of invariant measures for interval maps with positive Lyapunov exponent*, Stoch. Dyn. **9** (2009), no. 1, 81–100. MR MR2502475
8. N. Chernov and D. Kleinbock, *Dynamical Borel-Cantelli lemmas for Gibbs measures*, Israel J. Math. **122** (2001), 1–27. MR MR1826488 (2002h:37003)
9. P. Collet, *Statistics of closest return for some non-uniformly hyperbolic systems*, Ergodic Theory Dynam. Systems **21** (2001), no. 2, 401–420. MR MR1827111 (2002a:37038)
10. Dmitry Dolgopyat, *Limit theorems for partially hyperbolic systems*, Trans. Amer. Math. Soc. **356** (2004), no. 4, 1637–1689 (electronic). MR MR2034323 (2005k:37053)
11. Richard Durrett, *Probability: theory and examples*, second ed., Duxbury Press, Belmont, CA, 1996. MR MR1609153 (98m:60001)
12. Ana Cristina Moreira Freitas and Jorge Milhazes Freitas, *On the link between dependence and independence in extreme value theory for dynamical systems*, Statist. Probab. Lett. **78** (2008), no. 9, 1088–1093. MR MR2422964 (2009e:37006)

13. Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, and Mike Todd, *Hitting time statistics and extreme value theory*, Probability Theory and Related Fields (2009), published online.
14. Janos Galambos, *The asymptotic theory of extreme order statistics*, second ed., Robert E. Krieger Publishing Co. Inc., Melbourne, FL, 1987. MR MR936631 (89a:60059)
15. Sébastien Gouëzel, *A Borel-Cantelli lemma for intermittent interval maps*, Nonlinearity **20** (2007), no. 6, 1491–1497. MR MR2327135 (2008c:37012)
16. Chinmaya Gupta, *Extreme value distributions for some classes of non-uniformly partially hyperbolic dynamical systems*, to appear in Ergodic Theory and Dynamical Systems, 2009.
17. Chinmaya Gupta, Mark Holland, and Matthew Nicol, *Extreme value theory for hyperbolic billiards, Lozi-like maps, and Lorenz-like maps*, preprint, 2009.
18. Masaki Hirata, *Poisson law for Axiom A diffeomorphisms*, Ergodic Theory Dynam. Systems **13** (1993), no. 3, 533–556. MR MR1245828 (94m:58137)
19. Franz Hofbauer and Gerhard Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Math. Z. **180** (1982), no. 1, 119–140. MR MR656227 (83h:28028)
20. Mark Holland, Matthew Nicol, and Andrei Török, *Extreme value distributions for non-uniformly hyperbolic dynamical systems*, preprint.
21. Dong Han Kim, *The dynamical Borel-Cantelli lemma for interval maps*, Discrete Contin. Dyn. Syst. **17** (2007), no. 4, 891–900. MR MR2276480 (2007k:37008)
22. D. Y. Kleinbock and G. A. Margulis, *Logarithm laws for flows on homogeneous spaces*, Invent. Math. **138** (1999), no. 3, 451–494. MR MR1719827 (2001i:37046)
23. M. R. Leadbetter, Georg Lindgren, and Holger Rootzén, *Extremes and related properties of random sequences and processes*, Springer Series in Statistics, Springer-Verlag, New York, 1983. MR MR691492 (84h:60050)
24. Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory Dynam. Systems **19** (1999), no. 3, 671–685. MR MR1695915 (2000d:37029)
25. François Maucourant, *Dynamical Borel-Cantelli lemma for hyperbolic spaces*, Israel J. Math. **152** (2006), 143–155. MR MR2214457 (2008e:37028)
26. Ana Cristina Moreira Freitas and Jorge Milhazes Freitas, *Extreme values for Benedicks-Carleson quadratic maps*, Ergodic Theory Dynam. Systems **28** (2008), no. 4, 1117–1133. MR MR2437222
27. Y. Pomeau and P. Manneville, *Intermittent transition to turbulence in dissipative dynamical systems*, Commun. Math. Phys. **74** (1980), 189–197.
28. Sidney I. Resnick, *Extreme values, regular variation and point processes*, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2008, Reprint of the 1987 original. MR MR2364939 (2008h:60002)
29. Marek Rychlik, *Bounded variation and invariant measures*, Studia Math. **76** (1983), no. 1, 69–80. MR MR728198 (85h:28019)
30. Vladimir G. Sprindžuk, *Metric theory of Diophantine approximations*, V. H. Winston & Sons, Washington, D.C., 1979, Translated from the Russian and edited by Richard A. Silverman, With a foreword by Donald J. Newman, Scripta Series in Mathematics. MR MR548467 (80k:10048)
31. Lai-Sang Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650. MR MR1637655 (99h:58140)
32. ———, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188. MR MR1750438 (2001j:37062)

Department of Mathematics, 651 Philip G. Hoffman Hall, University of Houston, Houston, TX 77204
URL, Chinmaya Gupta: `http://www.math.uh.edu/~ccgupta`
E-mail address, Chinmaya Gupta: `ccgupta@math.uh.edu`

URL, Matthew Nicol: `http://www.math.uh.edu/~nicol`
E-mail address, Matthew Nicol: `nicol@math.uh.edu`

URL, William Ott: `http://www.math.uh.edu/~ott`