IOP Publishing | London Mathematical Society Nonlinearity

Nonlinearity **33** (2020) 892–914 https://doi.org/10.1088/1361-6544/ab5160

Shrinking targets and eventually always hitting points for interval maps

Maxim Kirsebom[1], Philipp Kunde[2] and Tomas Persson[3]

[1] Department of Mathematics, University of Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany
[2] Department of Mathematics, Pennsylvania State University, McAllister Building, State College, PA 16802, United States of America
[3] Centre for Mathematical Sciences, Lund University, Box 118, 221 00 Lund, Sweden

E-mail: maxim.kirsebom@uni-hamburg.de, pkunde.math@gmail.com and tomasp@maths.lth.se

Received 3 April 2019, revised 2 September 2019
Accepted for publication 25 October 2019
Published 10 January 2020

CrossMark

Recommended by Dr Sebastien Gouezel

Abstract
We study shrinking target problems and the set $\mathcal{E}_{\text{ah}}$ of eventually always hitting points. These are the points whose first n iterates will never have empty intersection with the nth target for sufficiently large n. We derive necessary and sufficient conditions on the shrinking rate of the targets for $\mathcal{E}_{\text{ah}}$ to be of full or zero measure especially for some interval maps including the doubling map, some quadratic maps and the Manneville–Pomeau map. We also obtain results for the Gauß map and correspondingly for the maximal digits in continued fraction expansions. In the case of β-transformations we also compute the packing dimension of $\mathcal{E}_{\text{ah}}$ complementing already known results on the Hausdorff dimension of $\mathcal{E}_{\text{ah}}$.

Keywords: shrinking target problems, eventually always hitting points, dynamical Borel–Cantelli, hitting time statistics, interval maps, continued fractions
Mathematics Subject Classification numbers: 37A05, 37E05, 11J70

1. Introduction and setup

The term *shrinking target problems* in dynamical systems describes a class of questions which seek to understand the recurrence behaviour of typical orbits of a dynamical system. The standard ingredients of such questions are a measure-preserving dynamical system (X, μ, T), $T: X \to X$ and μ being a finite measure. Also, we have a sequence of subsets $\{B_m\}_{m=1}^{\infty}$ with

$B_m \subset X$ and $\mu(B_m) \to 0$. Recently shrinking target problems have also been investigated in the case when the measure is infinite, see [11].

In this paper we focus mostly on the case of finite measure and whenever this is the case we assume the measure to be normalized to a probability measure. If nothing else is stated this may be assumed to be the setting. A few of our results concern infinite measures and it will be stated explicitly whenever this is the case.

Throughout this paper (X, μ, T) will always denote a measure preserving system and $\mathbf{B} := \{B_m\}_{m=1}^{\infty}$ will always denote a sequence of subsets of X for which $\mu(B_m) \to 0$. We refer to this as a *sequence of shrinking targets*. We call the sequence *nested* if $B_m \supset B_{m+1}$ for all m.

Classical questions in this area focus on the set of points in X whose nth iterate under T lies in the set B_n for infinitely many n. That is, given a sequence $\mathbf{B} = \{B_m\}_{m=1}^{\infty}$

$$\mathcal{A}_{\text{i.o.}} = \mathcal{A}_{\text{i.o.}}(\mathbf{B}) := \{x \in X : T^n x \in B_n \text{ for infinitely many } n \in \mathbb{N}\}.$$

If $\sum \mu(B_m) < \infty$ the Borel–Cantelli lemma tells us that $\mu(\mathcal{A}_{\text{i.o.}}) = 0$. If $\sum \mu(B_m) = \infty$ the situation is more complicated since the Borel–Cantelli lemma only guarantees $\mu(\mathcal{A}_{\text{i.o.}}) = 1$ for independent events and this is usually not satisfied for dynamical systems. If we do have a sequence $\mathbf{B}$ for which $\mu(\mathcal{A}_{\text{i.o.}}(\mathbf{B})) = 1$ then we call $\mathbf{B}$ a *Borel–Cantelli (BC) sequence*. If we can prove that a large family $\mathcal{B}$ of sequences are all BC-sequences then we say that we have a *dynamical Borel–Cantelli lemma*. In many cases such lemmas hold if the system satisfies some version of mixing which essentially acts as a replacement for independence. However, it is known that for any measure-preserving system we can find a sequence $\mathbf{B}$ satisfying $\sum \mu(B_m) = \infty$ which is not BC for the system. We may even find such a sequence $\mathbf{B}$ which is nested [4, proposition 1.6]. Hence there is no hope for $\mathcal{B}$ to be all sequences satisfying $\sum \mu(B_m) = \infty$. It is therefore natural to look for the largest possible sub-families of $\mathcal{B}$ which consist only of BC-sequences. It turns out that on metric spaces sequences of balls with fixed center, and nested sequences of balls with fixed center are good and natural candidates. We say that (X, μ, T) has the *shrinking target property* (STP) if for every $x_0 \in X$, every sequence of balls B_m centered at x_0 satisfying $\sum \mu(B_m) = \infty$ is a BC-sequence. We say that (X, μ, T) has the *monotone shrinking target property* (MSTP) if for every $x_0 \in X$, every nested sequence of balls B_m centered at x_0 satisfying $\sum \mu(B_m) = \infty$ is a BC-sequence. Many interesting systems are known to have either the STP or MSTP property, see [2, 25], and references therein for examples. A more comprehensive introduction to dynamical Borel–Cantelli lemmas, including examples, can be found in [4].

In this paper we are interested in similar properties for a certain subset of $\mathcal{A}_{\text{i.o.}}$ known as the set of points which are *eventually always hitting*. (More precisely, we consider a subset of $\mathcal{A}_{\text{i.o.}} \cup \Lambda$, where Λ is a set of zero measure, see (1) for details.) Due to the central importance of this concept in this paper we introduce it through a separate definition.

Definition 1 (Eventually always hitting). A point $x \in X$ is said to be *eventually always hitting* (EAH) for $\mathbf{B} = \{B_m\}_{m=1}^{\infty}$ under T if there exists some $m_0(x) \in \mathbb{N}$ such that for all $m \geqslant m_0(x)$ we have

$$\left\{x, T(x), T^2(x), \ldots, T^{m-1}(x)\right\} \cap B_m \neq \emptyset.$$

The set of all points in X which are eventually always hitting for $\mathbf{B}$ under T will be denoted $\mathcal{E}_{\text{ah}} := \mathcal{E}_{\text{ah}}(\mathbf{B})$.

We remark that some authors study a slightly different version of eventually almost hitting, and require that for all $m \geqslant m_0(x)$ we have $T^k(x) \in B_m$ for some k with $1 \leqslant k \leqslant m$, whereas we require $0 \leqslant k < m$. For the results that we are discussing in this paper, it is insignificant which definition we use. The results are the same, with the same proofs, if we use the other definition instead.

Note that $\mathcal{E}_{\mathrm{ah}}$ may also be written as

$$\mathcal{E}_{\mathrm{ah}} = \left\{ x \in X : \begin{array}{l} \exists\, m_0(x) \in \mathbb{N}\ \forall\, m > m_0(x) \\ \exists\, k \in \{0, \ldots, m-1\}\ \text{s.t.}\ T^k(x) \in B_m \end{array} \right\}$$
$$= \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \bigcup_{k=0}^{m-1} T^{-k}(B_m).$$

From this point on we will always assume the sequence $\{B_m\}$ to be nested. Set

$$\Lambda := \bigcup_{k=0}^{\infty} T^{-k}\left(\bigcap_{i=1}^{\infty} B_i\right). \tag{1}$$

Then $\mu(\Lambda) = 0$ since $\mu(B_i) \to 0$, and we have that $\mathcal{E}_{\mathrm{ah}} \backslash \Lambda \subset \mathcal{A}_{\mathrm{i.o.}}$. In this sense, being eventually always hitting for B_m is a stronger property than hitting B_m infinitely often.

The term eventually always hitting was coined by Kelmer in [14] where this set was studied in the context of flows on hyperbolic manifolds. Kelmer proved necessary and sufficient conditions for the set of eventually always hitting points to be of full measure. Shortly afterwards Kelmer and Yu [15] extended the investigation to flows on homogeneous spaces. Also, Kleinbock and Wadleigh [17] studied the concept in the context of higher dimensional Diophantine approximations and Oh and Kelmer considered the case of geodesic flow on geometrically finite hyperbolic manifolds [21]. Imposing a long-term independence property on the shrinking target system Kleinbock *et al* [16] recently obtained tight conditions on the shrinking rate of the targets so that $\mathcal{E}_{\mathrm{ah}}$ has measure zero or full measure. In particular, their assumptions are satisfied for specific choices of targets in product systems and Bernoulli shifts.

However, the concept had already been considered a few years earlier by Bugeaud and Liao [3] for a particular sequence of targets with exponential rate of shrinking in the setting of β-transformations $T_\beta(x) = \beta x \bmod 1$ on $[0,1]$ for every $\beta > 1$. For $x \in [0,1]$ they introduce the exponent $\hat{\nu}_\beta(x)$ as the supremum of real numbers $\hat{\nu}$ for which for every sufficiently large $N \in \mathbb{N}$ the inequality $T_\beta^n(x) < \left(\beta^N\right)^{-\hat{\nu}}$ has a solution $1 \leqslant n \leqslant N$. Note that this corresponds to x satisfying $\{T_\beta^n(x)\}_{n=1}^N \cap [0, \beta^{-\hat{\nu}N}) \neq \emptyset$ for every sufficiently large N. Hence the set $\{x \in [0,1] : \hat{\nu}_\beta(x) \geqslant \hat{\nu}\}$ corresponds to $\mathcal{E}_{\mathrm{ah}}([0, \beta^{-\hat{\nu}N}))$ in our notation (aside from the discrepancy in definition mentioned above). They show that

$$\dim_{\mathrm{H}}\left(\{\, x : \hat{\nu}_\beta(x) \geqslant \hat{\nu}\,\}\right) = \left(\frac{1-\hat{\nu}}{1+\hat{\nu}}\right)^2$$

for every $\hat{\nu} \in [0,1]$, where $\dim_{\mathrm{H}}$ is the Hausdorff dimension. Bugeaud and Liao also obtain analogous results in the setting of b-ary expansions.

In this paper we prove various results concerning the measure and also dimension of $\mathcal{E}_{\mathrm{ah}}$ in different settings. Our results apply in various levels of generality, hence it would be complicated to state them all accurately in this introduction. Instead we illustrate our results through application to specific systems with simpler assumptions.

Note that throughout the paper log will denote the natural logarithm.

1.1. Main results for specific systems

Theorem 1 (The doubling map). *Let $X=[0,1]$, let $T\colon X\to X$ be the doubling map $Tx=2x \bmod 1$ and let μ denote the Lebesgue measure on X. Let $\{B_m\}_{m=1}^{\infty}$ denote a nested sequence of shrinking intervals with fixed center.*

(a) If $\mu(B_m)\leqslant \frac{c}{m}$ for some $c\in\mathbb{R}$, then $\mu(\mathcal{E}_{\mathrm{ah}})=0$.
(b) If $\mu(B_m)\geqslant \frac{c(\log m)^2}{m}$ for some $c>0$ sufficiently large, then $\mu(\mathcal{E}_{\mathrm{ah}})=1$.

We remark that through our corollary 2 and theorem 5 we prove this theorem for many other dynamical systems, for instance piecewise expanding maps and some quadratic maps. See section 4 for more details.

We note that for the doubling map, $B_m=B(0,2^{-sm})$, and $0<s<1$, the result of Bugeaud and Liao [3] implies that $\dim_{\mathrm{H}}\mathcal{E}_{\mathrm{ah}}=\left(\frac{1-s}{1+s}\right)^2$. Hence, for the doubling map we have an almost complete picture of how the the size of $\mathcal{E}_{\mathrm{ah}}$ behaves across the spectrum of possible shrinking rates of B_m. In the following theorem we add to this picture by computing the packing dimension of $\mathcal{E}_{\mathrm{ah}}$.

Theorem 2. *Let $X=[0,1]$, let $T\colon X\to X$ be the β-transformation for some $\beta>1$, $Tx=\beta x \bmod 1$. Let $B_m=B(0,\beta^{-sm})$, $s>0$ and let $\dim_{\mathrm{P}}$ denote the packing dimension.*

(a) If $s\in(0,1)$, then $\dim_{\mathrm{P}}\mathcal{E}_{\mathrm{ah}}=1-s$.
(b) If $s\geqslant 1$ then $\mathcal{E}_{\mathrm{ah}}=\Lambda$. Hence it is countable and in particular $\dim_{\mathrm{P}}\mathcal{E}_{\mathrm{ah}}=0$.

Figure 1 illustrates and compares the results of theorems 1 and 2 for the doubling map. It summarises the known results on the measure, Hausdorff dimension and packing dimension of the sets $\mathcal{A}_{\mathrm{i.o.}}$ and $\mathcal{E}_{\mathrm{ah}}$. Note that if $\mu(B_m)=2^{-sm}$, then $\dim_{\mathrm{H}}\mathcal{A}_{\mathrm{i.o.}}=\frac{1}{1+s}$ [1, corollary 1].

Theorem 3 (The Manneville–Pomeau map). *Let $X=[0,1]$, let $\alpha>0$ and let $g_\alpha\colon[0,1)\to[0,1)$ be the Manneville–Pomeau map given by*

$$g_\alpha(x)=\begin{cases}x(1+2^\alpha x^\alpha) & \text{if } x\in[0,\frac{1}{2})\\ 2x-1 & \text{if } x\in[\frac{1}{2},1).\end{cases}\tag{2}$$

Let μ_α denote the absolutely continuous to the Lebesgue measure invariant measure for $([0,1),g_\alpha)$, which is finite if and only if $\alpha\in(0,1)$. Let $\{B_m\}_{m=1}^{\infty}$ denote a nested sequence of intervals with a fixed center which is not 0.

If $\alpha\in(0,1)$ then we have the following results.

(a) If $\mu_\alpha(B_m)\leqslant\frac{c}{m}$ for some constant c, then $\mu_\alpha(\mathcal{E}_{\mathrm{ah}})=0$.
(b) If $\mu_\alpha(B_m)\geqslant\frac{c(\log m)^{2+\varepsilon}}{m}$ for some $\varepsilon>0$ and any $c>0$, then we have $\mu_\alpha(\mathcal{E}_{\mathrm{ah}})=1$.

If $\alpha\geqslant1$ then we have the following results.

(c) If $\alpha>1$ and $\mu_\alpha(B_m)\leqslant\frac{c}{m^{\frac{1}{\alpha}+\varepsilon}}$ for some $\varepsilon>0$ and any $c>0$, then $\mu_\alpha(\mathcal{E}_{\mathrm{ah}})=0$.
If $\alpha=1$ and $\mu_\alpha(B_m)\leqslant\frac{c}{m}$ for some $c>0$ then $\mu_\alpha(\mathcal{E}_{\mathrm{ah}})=0$.
(d) If $\mu_\alpha(B_m)\geqslant\frac{c}{m^{\frac{1}{\alpha}-\varepsilon}}$ for some $\varepsilon>0$ and any $c>0$, then $\mu_\alpha(\complement\mathcal{E}_{\mathrm{ah}})=0$[4].

[4] Throughout the paper we denote the complement of a set A by $\complement A$.

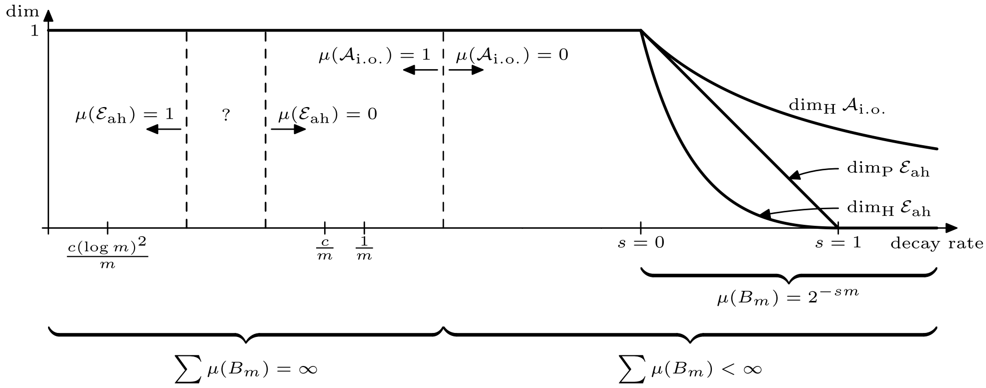

Figure 1. Illustration of the sizes of the sets $\mathcal{A}_{\text{i.o.}}$ and $\mathcal{E}_{\text{ah}}$ for different decay rates of $\mu(B_m)$, in the case of the doubling map.

We remark that the measure μ_α is equivalent to Lebesgue measure on $[0, 1]$. Hence the statements $\mu_\alpha(\mathcal{E}_{\text{ah}}) = 0$ and $\mu_\alpha(\complement\mathcal{E}_{\text{ah}}) = 0$ are equivalent to the corresponding statements involving the Lebesgue measure instead. Note also that in the case $\alpha \in (0, 1)$ after normalising the measure μ_α we can state $\mu_\alpha(\complement\mathcal{E}_{\text{ah}}) = 0$ as the equivalent statement $\mu_\alpha(\mathcal{E}_{\text{ah}}) = 1$. This is not possible if $\alpha \geqslant 1$, since the measure μ_α is not finite in this case.

We may also consider the Gauß map defined by $G : (0, 1] \to [0, 1)$ given by $G(x) = \frac{1}{x} \bmod 1$. The map G admits an absolutely continuous invariant probability measure known as the Gauß measure which has density $\frac{1}{\log 2}\frac{1}{1+x}$. Statement (b) of theorem 1 also holds true for the Gauß map and measure, while (a) of theorem 1 holds true in this setting when $B_m := [0, r_m)$ with $r_m \leqslant \frac{c}{m}$. This allows us to obtain a statement about the eventually always hitting property for maximal digits of continued fractions expansions. Recall that every point $x \in (0, 1]$ can be written as a continued fraction, i.e.

$$x = \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \cfrac{1}{a_3(x) + \cfrac{1}{\ddots}}}}$$

where the $a_i(x)$'s are generated by the algorithm $a_1(x) = \lfloor \frac{1}{x} \rfloor$ and $a_j(x) = a_1(G^{j-1}(x))$. In compact notation we write $x = [a_1(x), a_2(x), \ldots]$.

Theorem 4 (Continued fractions). *Let μ denote the Gauß measure.*

(a) For any $c > 0$ we have

$$\mu\left(\left\{ x \in [0, 1] : \exists m_0 \in \mathbb{N}\, \forall m \geqslant m_0 : \max_{1 \leqslant k \leqslant m} a_k(x) \geqslant cm \right\}\right) = 0.$$

(b) If $c > 0$ is sufficiently small, then

$$\mu\left(\left\{ x \in [0, 1] : \exists m_0 \in \mathbb{N}\, \forall m \geqslant m_0 : \max_{1 \leqslant k \leqslant m} a_k(x) \geqslant \frac{cm}{(\log m)^2} \right\}\right) = 1.$$

1.2. Structure of the paper

In section 2 we give some preliminary, general results concerning the set of eventually always hitting points that will prove useful later in the paper. In section 3 we give necessary and sufficient conditions for $\mathcal{E}_{\mathrm{ah}}$ to be of full measure. In section 4 we cite various results on mixing and hitting time statistics which in conjunction with the results of section 3 allow us to deduce the conclusions of theorems 1, 3 and 4 (except theorems 1(a) and (b)). We also discuss further systems for which the results of section 3 can be applied. Section 5 is dedicated to the proof of theorems 1(a) and (b).

2. Preliminaries on eventually always hitting points

Recall that

$$\mathcal{E}_{\mathrm{ah}} = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \bigcup_{k=0}^{m-1} T^{-k}(B_m).$$

It will prove convenient to write

$$\mathcal{E}_{\mathrm{ah}} = \bigcup_{n=1}^{\infty} A_n$$

where

$$A_n = \bigcap_{m=n}^{\infty} C_m \quad , \quad C_m = \bigcup_{k=0}^{m-1} T^{-k}(B_m).$$

Note in particular that $A_n \subset A_{n+1}$ and hence

$$\mu(\mathcal{E}_{\mathrm{ah}}) = \lim_{n\to\infty} \mu(A_n).$$

The following lemma will prove very useful for studying the measure of $\mathcal{E}_{\mathrm{ah}}$ when the sets B_m are assumed to be nested. It states that $\mathcal{E}_{\mathrm{ah}}$ is an essentially invariant set under T. More precisely, the symmetric difference between $\mathcal{E}_{\mathrm{ah}}$ and $T^{-1}(\mathcal{E}_{\mathrm{ah}})$ is a set of measure zero.

Lemma 1. *Let (X, μ, T) be a measure-preserving dynamical system (μ either finite or infinite) and let $\{B_m\}_{m=1}^{\infty}$ be a nested family of shrinking targets, i.e. $B_m \supset B_{m+1}$ and $\mu(B_m) \to 0$, with $\mu(\mathcal{E}) < \infty$. Then*

$$\mu\left(\mathcal{E}_{\mathrm{ah}} \triangle T^{-1}\left(\mathcal{E}_{\mathrm{ah}}\right)\right) = 0.$$

Hence, for ergodic transformations with respect to a finite μ, $\mathcal{E}_{\mathrm{ah}}$ obeys a zero–one law. That is, either $\mu\left(\mathcal{E}_{\mathrm{ah}}\right) = 0$ or $\mu\left(\mathcal{E}_{\mathrm{ah}}\right) = 1$. For ergodic transformations with respect to an infinite μ, $\mathcal{E}_{\mathrm{ah}}$ obeys a zero–infinity law.

Proof. Let $x \in \mathcal{E}_{\mathrm{ah}}$. Then there is $m_0 := m_0(x) \in \mathbb{N}$ such that for all $m \geqslant m_0$ there is $k \in \{0, 1, \ldots, m-1\}$ with $T^k(x) \in B_m$. Actually, if $x \notin \bigcap_{m\in\mathbb{N}} B_m$ there is even $\tilde{m}_0 := \tilde{m}_0(x) \in \mathbb{N}$ such that for all $m \geqslant \tilde{m}_0$ there is $k \in \{1, \ldots, m-1\}$ with $T^k(x) \in B_m$ (because otherwise $x = T^0(x)$ would have to be in B_m for all m due to $x \in \mathcal{E}_{\mathrm{ah}}$ and the nesting property). Since the targets are nested, we also get for those $x \in \mathcal{E}_{\mathrm{ah}} \setminus \bigcap_{m\in\mathbb{N}} B_m$ that for all $m \geqslant \tilde{m}_0$ there is $k \in \{1, \ldots, m-1\}$ with

$$T^{k-1}(T(x)) = T^k(x) \in B_m \subseteq B_{m-1}.$$

Hence, $T(x) \in \mathcal{E}_{\mathrm{ah}}$. So, $\mathcal{E}_{\mathrm{ah}} \setminus \bigcap_{m\in\mathbb{N}} B_m \subseteq T^{-1}(\mathcal{E}_{\mathrm{ah}})$. Since $\mu\left(\bigcap_{m\in\mathbb{N}} B_m\right)$ is a set of measure zero by $\mu(B_m) \to 0$ and T is measure-preserving, it follows that

$$\mu(\mathcal{E}_{\mathrm{ah}}) = \mu\left(\mathcal{E}_{\mathrm{ah}} \setminus \bigcap_{m\in\mathbb{N}} B_m\right) \leqslant \mu\left(T^{-1}(\mathcal{E}_{\mathrm{ah}})\right) = \mu(\mathcal{E}_{\mathrm{ah}}),$$

which implies that $\mathcal{E}_{\mathrm{ah}} \setminus \bigcap_{m\in\mathbb{N}} B_m$ and $T^{-1}(\mathcal{E}_{\mathrm{ah}})$ differ by a null set. Since $\mu\left(\bigcap_{m\in\mathbb{N}} B_m\right) = 0$, the claim follows. □

In [14] Kelmer gave the following simple conditions for $\mathcal{E}_{\mathrm{ah}}$ to be of measure zero or one. We repeat the proof for completeness. Note that no assumption is made on the shape of the target sets.

Proposition 1. *Let $\{B_m\}_{m=1}^{\infty}$ denote a sequence of shrinking targets in X.*

(a) Let μ be a probability measure and assume that $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$. Then there exists a sequence $c_m \to 1$ such that $\mu(B_m) \geqslant \frac{c_m}{m}$.
(b) Let μ be an infinite measure and assume that $\mu(\mathcal{E}_{\mathrm{ah}}) = \infty$. Then there exists a sequence $c_m \to \infty$ such that $\mu(B_m) \geqslant \frac{c_m}{m}$.

Proof. (a) By the assumption we get that

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \mu\left(\bigcap_{m=n}^{\infty} C_m\right) = 1$$

which implies that $\mu(C_m) \to 1$ as $m \to \infty$. Now,

$$\mu(C_m) = \mu\left(\bigcup_{k=0}^{m-1} T^{-k}(B_m)\right) \leqslant \sum_{k=0}^{m-1} \mu\big(T^{-k}(B_m)\big) = m\mu(B_m)$$

where we used the T-invariance of μ. Hence we have $\mu(B_m) \geqslant \frac{\mu(C_m)}{m}$. Since $\mu(C_m) \to 1$, we obtain (a).
(b) The proof goes exactly like for (a) with the obvious adaptations. □

For nested sequences $\{B_m\}_{m=1}^{\infty}$ and T being ergodic we get the following sufficient condition for $\mu(\mathcal{E}_{\mathrm{ah}}) = 0$. The proof is just the negation of proposition 1 followed by an application of lemma 1.

Corollary 1. *Assume that T is ergodic and that $\{B_m\}_{m=1}^{\infty}$ is a nested sequence of shrinking targets.*

(a) Let μ be a probability measure. If there exists a $c < 1$ such that $\mu(B_m) \leqslant \frac{c}{m}$ holds for infinitely many m, then $\mu(\mathcal{E}_{\mathrm{ah}}) = 0$.
(b) Let μ be an infinite measure. If there exists a $c \in \mathbb{R}$ such that $\mu(B_m) \leqslant \frac{c}{m}$ holds for infinitely many m, then $\mu(\mathcal{E}_{\mathrm{ah}}) = 0$.

3. Necessary and sufficient conditions for $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$

In this section we give proofs of various new necessary and sufficient conditions for $\mu(\mathcal{E}_{\mathrm{ah}})$ to be of measure zero or one.

3.1. Necessary conditions for $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$

We introduce some terminology and notation about hitting times in dynamical system. Given a set $E \subset X$, we denote by $\tau_E \colon X \to \mathbb{N}$ the *first hitting time to* E which is defined by

$$\tau_E(x) = \inf\left\{i \in \mathbb{N} : T^i x \in E\right\}.$$

Let $E_m \subset X$ denote a sequence of sets for which $\mu(E_m) \to 0$ and define the function

$$G_{E_m}(t) := \limsup_{m\to\infty} \mu\left(\left\{x \in X : \tau_{E_m}(x) \leqslant \frac{t}{\mu(E_m)}\right\}\right).$$

The next easy proposition gives a necessary condition for $\mathcal{E}_{\mathrm{ah}}$ to be of full measure when $G(t) < 1$ for all $t \in \mathbb{R}$. We note that while this condition might appear arbitrary at this point, it is very often satisfied for dynamical systems. Indeed, it is a weaker condition than the system having exponential hitting time statistics, a concept much studied and often satisfied in dynamics. Later in this section we discuss hitting time statistics and in section 4 we discuss examples of systems where this property is known.

Proposition 2. *Let* $\{B_m\}_{m=1}^{\infty}$ *denote a sequence of shrinking targets and assume that* $G_{B_m}(t) < 1$ *for all* $t \in \mathbb{R}$*. If* $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$*, then for every* $c \in \mathbb{R}$ *we have* $\mu(B_m) \geqslant \frac{c}{m}$ *for all sufficiently large* $m \in \mathbb{N}$*.*

Proof. To get a contradiction, fix $c \in \mathbb{R}$, and assume that there is a sequence $m_j \to \infty$ such that $\mu(B_{m_j}) \leqslant \frac{c}{m_j}$ for all j. We then have

$$\left\{x \in X : \tau_{B_{m_j}}(x) \leqslant m_j\right\} \subset \left\{x \in X : \tau_{B_{m_j}}(x) \leqslant \frac{c}{\mu(B_{m_j})}\right\}.$$

Using this inclusion we may rewrite as follows

$$\begin{aligned}
\mathcal{E}_{\mathrm{ah}} &= \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{x \in X : \left\{T^i x\right\}_{i=0}^{m-1} \cap B_m \neq \emptyset\right\} \\
&= \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \{x \in X : \tau_{B_m}(x) < m\} \\
&\subset \bigcup_{n=1}^{\infty} \bigcap_{m_j \geqslant n} \left\{x \in X : \tau_{B_{m_j}}(x) \leqslant m_j\right\} \subset \bigcup_{n=1}^{\infty} \bigcap_{m_j \geqslant n} \tilde{C}_{m_j},
\end{aligned}$$

where

$$\tilde{C}_{m_j} = \left\{x \in X : \tau_{B_{m_j}}(x) \leqslant \frac{c}{\mu(B_{m_j})}\right\}.$$

Assuming that $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$, we argue as in proposition 1 that we must have $\mu(\tilde{C}_{m_j}) \to 1$ which means that

$$\mu\left(\left\{x \in X : \tau_{B_{m_j}}(x) \leqslant \frac{c}{\mu(B_{m_j})}\right\}\right) \to 1$$

for $m_j \to \infty$. However, $\limsup_{j\to\infty} \mu(\tilde{C}_{m_j}) \leqslant G_{B_m}(c) < 1$ by assumption and hence we have a contradiction. Since this inequality is true for all $c \in \mathbb{R}$ we get the desired conclusion. $\square$

Again we get a sufficient condition for $\mu(\mathcal{E}_{\mathrm{ah}}) = 0$ under additional assumptions.

Corollary 2. *Assume that T is ergodic and that $\{B_m\}_{m=1}^{\infty}$ is a nested sequence of shrinking targets and assume that $G_{B_m}(t) < 1$ for all $t \in \mathbb{R}$. If there exists $c \in \mathbb{R}$ such that $\mu(B_m) \leqslant \frac{c}{m}$ for infinitely many m, then $\mu(\mathcal{E}_{\mathrm{ah}}) = 0$.*

As already mentioned, the condition $G_{B_m}(t) < 1$ is easily satisfied in many cases. It is interesting to note that often much more is known about $G_{B_m}(t)$. If we assume B_m to be a nested sequence of shrinking balls with fixed center it is often known that

$$G_{B_m}(t) = \lim_{m\to\infty} \mu\left(\left\{x \in X : \tau_{B_m}(x) \leqslant \frac{t}{\mu(B_m)}\right\}\right)$$

exists and is non-degenerate which means that $G_{B_m}(t)$ takes at least one value different than 0 or 1. This property is known as the system having *hitting time statistics* (HTS) to B_m. Among these systems, many have exponential HTS to B_m meaning that $G_{B_m}(t) = 1 - \mathrm{e}^{-t}$. See [20, chapter 5] for a long list of examples of such systems. It is from the rich theory of HTS for dynamical systems that we borrow in order to prove theorems 1(a), 3(a) and 4(a). We elaborate on this point and give examples as well as exact statements statements concerning HTS in section 4.

3.2. Sufficient conditions for $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$

Let $X \subset \mathbb{R}$ in our probability measure preserving system (X, T, μ) and let $B_m = B(y_m, r_m)$ be a sequence of balls in X. We consider the L^1 and BV norms of functions on $f\colon X \to \mathbb{R}$, defined by

$$||f||_1 = \int |f|\,\mathrm{d}\mu,$$
$$||f||_{BV} = \mathrm{var} f + ||f||_1,$$

where $\mathrm{var} f$ denotes the total variation of f.

We say that correlations decay as $p\colon \mathbb{N} \to \mathbb{R}$ for L^1 against BV, if

$$\left|\int f \circ T^n g\,\mathrm{d}\mu - \int f\,\mathrm{d}\mu \int g\,\mathrm{d}\mu\right| \leqslant ||f||_1 ||g||_{BV} p(n) \tag{3}$$

holds for all n and all functions f and g with $||f||_1, ||g||_{BV} < \infty$.

Theorem 5. *Suppose that correlations decay as p for L^1 against BV.*

If p satisfies

$$p(n) \leqslant C\mathrm{e}^{-\tau n} \tag{4}$$

for some $\tau > 0$, then $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$ provided

$$\mu(B_m) \geqslant \frac{c(\log m)^2}{m}$$

for some $c > \tau^{-1}$ and all sufficiently large m. In particular $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$ provided $\mu(B_m) \geqslant \frac{c(\log m)^2 h(m)}{m}$ for any $c > 0$ and any function h for which $h(m) \to \infty$ as $m \to \infty$.

If p satisfies

$$p(n) \leqslant \frac{C}{n^t} \tag{5}$$

for some $t > 0$*, then* $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$ *provided*

$$\mu(B_m) \geqslant \frac{c}{m^a}$$

for any $c > 0$ *and* $a < \frac{t}{1+t}$.

Note that we do not require that the balls are nested in theorem 5. We will need the following lemma.

Lemma 2. *Suppose that (3) holds for all n and functions f and g. If* $f_k\colon X \to [0, \infty)$ *and* $n_k \in \mathbb{N}$*, then*

$$\int \prod_{k=1}^{n} f_k \circ T^{n_1+\cdots+n_k} \,\mathrm{d}\mu \leqslant \int f_n \,\mathrm{d}\mu \prod_{k=1}^{n-1} \left(\int f_k \,\mathrm{d}\mu + p(n_{k+1}) ||f_k||_{BV} \right).$$

Proof. Note that since $f_k \geqslant 0$, we have $\int f_k \,\mathrm{d}\mu = ||f_k||_1$. By (3), we have

$$\int \prod_{k=1}^{n} f_k \circ T^{n_1+\cdots+n_k} \,\mathrm{d}\mu \leqslant \left(\int f_1 \mathrm{d}\mu + p(n_2) ||f_1||_{BV} \right) \int \prod_{k=2}^{n} f_k \circ T^{n_1+\cdots+n_k} \,\mathrm{d}\mu,$$

and the inequality follows by induction. □

Proof of theorem 5. Recall that

$$\mathcal{E}_{\mathrm{ah}} = \bigcup_{n=1}^{\infty} A_n.$$

We prove that $\mu(A_n) \to 1$ as $n \to \infty$ which is equivalent to $\mu(\complement A_n) \to 0$. We can write $\complement A_n$ as

$$\complement A_n = \bigcup_{m=n}^{\infty} D_m, \qquad \text{where } D_m = \bigcap_{k=0}^{m-1} T^{-k}(\complement B_m).$$

Since $\mu\left(\complement A_n\right) \leqslant \sum_{m=n}^{\infty} \mu\left(D_m\right)$, it is sufficient to prove that

$$\sum_{m=1}^{\infty} \mu(D_m) < \infty. \tag{6}$$

We will now bound $\mu(D_m)$ from above. Let $\Delta_m > 1$[5]. We have that

$$\mu(D_m) \leqslant \mu(\tilde{D}_m), \qquad \text{where } \tilde{D}_m = \bigcap_{0 \leqslant k \leqslant (m/\Delta_m)-1} T^{-\Delta_m k}(\complement B_m).$$

Then, we have by lemma 2 that

[5] Since Δ_m is typically non-integer we should, in principle, be more diligent and write $\lfloor \Delta_m \rfloor$ and $\lfloor \Delta_m k \rfloor$ in the subsequent estimates. However, to improve readability we let the relevant ceiling and floor functions be implicitly understood throughout the proof. The outcome is invariant under this abuse of notation.

$$\mu(\tilde{D}_m) = \int \prod_{0 \leqslant k \leqslant (m/\Delta_m)-1} 1_{\complement B_m} \circ T^{\Delta_m k} \, \mathrm{d}\mu$$
$$\leqslant \prod_{0 \leqslant k \leqslant (m/\Delta_m)-1} \left(\mu(\complement B_m) + p(\Delta_m) ||1_{\complement B_m}||_{BV} \right).$$

Since B_m are balls, we have $\mathrm{var} 1_{\complement B_m} \leqslant 2$ and $||1_{\complement B_m}||_1 \leqslant 1$. Hence we have $||1_{\complement B_m}||_{BV} \leqslant 3$ and we get

$$\mu(\tilde{D}_m) \leqslant \prod_{0 \leqslant k \leqslant (m/\Delta_m)-1} \left(1 - \mu(B_m) + 3p(\Delta_m) \right)$$
$$= \exp\left(\sum_{0 \leqslant k \leqslant (m/\Delta_m)-1} \log\big(1 - \mu(B_m) + 3p(\Delta_m)\big) \right)$$
$$\leqslant \exp\left(\frac{m}{\Delta_m} 3p(\Delta_m) - \sum_{0 \leqslant k \leqslant (m/\Delta_m)-1} \mu(B_m) \right),$$

where the last inequality holds since $\log(1+x) \leqslant x$.

Assume now that $p(n) \leqslant C\mathrm{e}^{-\tau n}$. We prove that $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$ if for all sufficiently large m, we have $\mu(B_m) \geqslant \frac{c(\log m)^2}{m}$ where $c > \tau^{-1}$.

Take $\Delta_m = \frac{1}{\tau} \log m$. Then

$$\frac{m}{\Delta_m} 3p(\Delta_m) \to 0 \qquad \text{as } m \to \infty,$$

so we can assume $(m/\Delta_m) 3p(\Delta_m) < 1$ if m is large.

We assume that we have $\mu(B_m) \geqslant \frac{c(\log m)^2}{m}$ for all m. The proof also works with obvious changes if this is only the case for all large enough m. We have

$$\sum_{0 \leqslant k \leqslant (m/\Delta_m)-1} \mu(B_m) \geqslant \frac{m}{\Delta_m} \frac{c(\log m)^2}{m} = c\tau \log m.$$

Taken together, these two estimates give us the estimate

$$\mu(\tilde{D}_m) \leqslant \exp\big(1 - c\tau \log m\big),$$

which implies that $\mu(\tilde{D}_m)$ is summable since $c\tau > 1$.

We now consider the case that $p(n) \leqslant C/n^t$ and $\mu(B_m) \geqslant \frac{c}{m^a}$. Let $\Delta_m = m^{\frac{1}{1+t}}$. Then

$$\sum_{0 \leqslant k \leqslant (m/\Delta_m)-1} 3p(\Delta_m) \leqslant \frac{m}{\Delta_m} \frac{3C}{\Delta_m^t} = 3C.$$

Hence

$$\mu(\tilde{D}_m) \leqslant \exp\left(3C - \sum_{0 \leqslant k \leqslant m^{\frac{t}{1+t}}-1} \frac{c}{m^a} \right) = \exp\left(3C - cm^\beta \right),$$

where $\beta = \frac{t}{1+t} - a$. If $a < \frac{t}{1+t}$, then $\beta > 0$ and $\mu(\tilde{D}_m)$ is summable, which proves (6). □

We note that the sufficient condition obtained by Kelmer [14] in the setting of discrete time homogeneous flows acting on a finite volume quotient of $\mathbb{H}^n$ is slightly better than what we get in our setting. More precisely, [14, theorem 2] states that in the mentioned setting

$$\sum_{j=0}^{\infty} \frac{1}{2^j \mu(B_{2^j})} < \infty \qquad \Rightarrow \qquad \mu(\mathcal{E}_{\mathrm{ah}}) = 1.$$

Inserting either $\mu(B_m) \geqslant \frac{c(\log m)^2}{m}$ or $\mu(B_m) \geqslant \frac{c}{m^a}$ in the above both result in convergent sums. A bound like $\mu(B_m) \geqslant \frac{c(\log m)}{m}$ or lower would be a sharper bound compared to that of Kelmer. However, the settings are very different and it is not clear what an optimal bound would look like in either setting.

4. Application to examples

Many systems are known to have either polynomial or exponential decay of correlation for L^1 against BV in the sense of (4) and (5). Examples of exponential decay includes T being a piecewise expanding interval map and μ being a Gibbs measure, T being a quadratic map for a Benedicks–Carleson parameter and μ being the absolutely continuous invariant measure, [18, 19, 26]. Hence, for these systems, $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$ whenever $\mu(B_m) \geqslant \frac{c(\log m)^2}{m}$ for some $c > 0$ sufficiently large.

Hitting time statistics is known for many interesting dynamical systems. Since HTS is often true for systems with sufficiently nice mixing properties, our necessary and sufficient conditions tend to hold for many of the same systems. Examples of HTS in dynamics include transitive Markov chains, Axiom A diffeomorphisms, uniformly expanding maps of the interval, non-uniformly hyperbolic maps, partially hyperbolic dynamical systems and toral automorphisms. Hence, for these systems $\mu(\mathcal{E}_{\mathrm{ah}}) = 0$ whenever $\mu(B_m) \leqslant \frac{c}{m}$ for some $c \in \mathbb{R}$. For a comprehensive overview of these results and references, see [20, chapter 5]. Furthermore, [7, 8] establish a direct connection between HTS and so-called extreme value theory meaning that many HTS results can be obtained simply by translating known extreme value laws. See again [20] for an overview of such results.

In the following subsections we describe in more detail how theorems 1, 3 and 4 are deduced.

4.1. The doubling map

HTS is known to hold for the doubling map. The precise statement is as follows. Let $r_m \in \mathbb{R}$ be a sequence and set $E_m = B(y, r_m)$, i.e. the ball with center $y \in [0, 1]$ and radius r_m. For any sequence $r_m \to 0$ we have[6]

$$G_{E_m}(t) = \begin{cases} 1 - \mathrm{e}^{-t} & \text{if } y \text{ is not a periodic point} \\ 1 - \mathrm{e}^{-\left(1-\frac{1}{2}\right)^p t} & \text{if } y \text{ is a periodic point of prime period } p \in \mathbb{N}. \end{cases}$$

This explicit result can be deduced by applying [7, theorem 2] (along with [8, p 7]) to [20, corollary 4.2.11]. Hence theorem 1(a) follows directly from corollary 2.

[6] A point y is a periodic point with prime period p if p is the smallest natural number such that $T^p(y) = y$.

The doubling map with Lebesgue measure is a well-known example of a piecewise expanding interval map with a Gibbs measure. Hence it is exponentially mixing for L^1 against BV [19] and theorem 1(b) follows directly from theorem 5.

4.2. The Manneville–Pomeau map

Here we explain how theorem 3 follows from our results. We begin with theorem 3(a).

For the Manneville–Pomeau maps the following is known regarding hitting time statistics in the case when $\alpha < 1$, that is when the invariant measure μ_α is finite. Let $E_m = B(y, r_m)$ as above, then for any sequence $r_m \to 0$ we have

$$G_{E_m}(t) = \begin{cases} 1 - \mathrm{e}^{-t} & \text{if } y \text{ is not a periodic point} \\ 1 - \mathrm{e}^{-\left(1-\frac{1}{|Dg_\alpha^p(y)|}\right)t} & \text{if } y \text{ is a periodic point of prime period } p \\ 0 & \text{if } y = 0, \end{cases}$$

where $Dg_\alpha^p(y)$ denotes the derivative of g_α^p at the point y. This follows again by applying [7, theorem 2] (along with [8, p 7]) to [9, theorems 1 and 2]. Hence theorem 3(a) follows directly from corollary 2. Note that theorem 3(a) actually also holds for balls centered at 0.

We proceed to deducing theorems 3(b), (c) and (d). Note first that the case $\alpha = 1$ in theorem 3(c) follows directly from corollary 1(b).

The Manneville–Pomeau map is not known to have decay of correlations for L^1 against BV. However, through a technique known as *inducing*, explained below, we can obtain results almost as strong as if it had exponential decay of correlations with respect to said norms. We let $S\colon [\frac{1}{2}, 1) \to [\frac{1}{2}, 1)$ be the first return map of g_α to the interval $[\frac{1}{2}, 1)$. The structure of the map S is illustrated in figure 2.

Given a point $x \in [\frac{1}{2}, 1)$, there is then a sequence $R_k(x)$ such that

$$S^k(x) = g_\alpha^{R_k(x)}(x).$$

The sequence $R_k(x)$ satisfies

$$R_k(x) = \sum_{j=0}^{k-1} R(S^j(x)),$$

where R is the return time $R(x) = \min\{\, n \geqslant 1 : g_\alpha^n(x) \in [\frac{1}{2}, 1)\,\}$.

The return map S is uniformly expanding and it follows by the paper of Rychlik [23] that it has exponential decay of correlations for L^1 against BV. Hence we may apply theorem 5 to S.

The absolutely continuous invariant measure μ_α of g_α is finite on $[\frac{1}{2}, 1)$ and we write $\tilde{\mu}_\alpha$ for the normalized measure, i.e. $\tilde{\mu}_\alpha([\frac{1}{2}, 1)) = 1$. The measure μ_α is finite on $[0, 1)$ if and only if $\alpha \in (0, 1)$. In fact, the density h of the measure μ_α is a decreasing and positive function, and it satisfies $h(x) \sim x^{-\alpha}$ when x is close to zero [24].

Using the first return map enables us to estimate $\tilde{\mu}_\alpha(\mathcal{E}_{\mathrm{ah}}(\mathbf{B}) \cap [\frac{1}{2}, 1))$ for $\mathbf{B} = \{B(y, r_m)\}_{m=1}^{\infty}$ with center $y \in [\frac{1}{2}, 1)$ depending on the rate of shrinking of r_m. At the end of the section we argue why this is sufficient. For now set $B_m = B(y, r_m)$ and assume that $y \in [\frac{1}{2}, 1)$. By using the fact that $g_\alpha^n(x) \in B_m$ can only happen along the subsequence $n_k := R_k(x)$, a short argument gives the inclusions

$$A_1 \subset \mathcal{E}_{\mathrm{ah}} \cap \left[\frac{1}{2}, 1\right) \subset A_2 \tag{7}$$

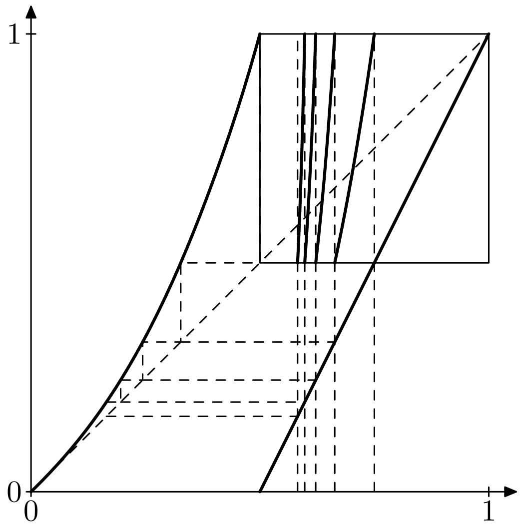

Figure 2. The Manneville–Pomeau map and its first return map.

where

$$A_1 = \left\{ x \in \left[\frac{1}{2}, 1\right) : \exists\, m_0\, \forall\, m \geqslant m_0\, \exists\, k < m : S^k(x) \in B_{R_{m+1}(x)} \right\},$$

and

$$A_2 = \left\{ x \in \left[\frac{1}{2}, 1\right) : \exists\, m_0\, \forall\, m \geqslant m_0\, \exists\, k < m : S^k(x) \in B_{R_m(x)} \right\}.$$

The small argument verifying (7) is left to the reader.

Assume first that μ_α is finite, i.e. $\alpha \in (0, 1)$. We will use (7) together with the following estimate of R_n from [11, theorem 2.19]. There exists a constant $C > 1$ such that the set of x for which the inequality,

$$n \leqslant R_n(x) \leqslant Cn \tag{8}$$

does not hold is of arbitrarily small measure if n is sufficiently large. Let N_0 be so large that the set of x for which (8) holds for $n > N_0$ is of measure at least $1 - \delta$ and call this set D_δ. For $x \in D_\delta$ we may apply (8) and we find that the condition $\tilde{\mu}_\alpha(B_m) \geqslant \frac{c(\log m)^{2+\varepsilon}}{m}$ implies that $\tilde{\mu}_\alpha(B_{R_n(x)}) \geqslant \frac{c(\log n)^{2+\varepsilon}}{Cn} > \frac{c_0(\log n)^2}{n}$ for large n. So by picking $C_m := B(y, r_m)$ with $\tilde{\mu}_\alpha(C_m) = \frac{c(\log m)^{2+\varepsilon}}{m}$ we get $\tilde{A}_1 \subset A_1$ where

$$\tilde{A}_1 := \left\{ x \in \left[\frac{1}{2}, 1\right) : \exists\, m_0\, \forall\, m \geqslant m_0\, \exists\, k < m : S^k(x) \in C_m \right\} \cap D_\delta.$$

The first set in the intersection is $\mathcal{E}_{\text{ah}}(C_m)$ for the system $([\frac{1}{2}, 1), S, \tilde{\mu}_\alpha)$ which has measure one by theorem 5. This shows that $\tilde{\mu}_\alpha(\tilde{A}_1) \geqslant 1 - \delta$ and since δ is arbitrary we get $\tilde{\mu}_\alpha(A_1) = 1$. This implies that $\tilde{\mu}_\alpha(\mathcal{E}_{\text{ah}} \cap [\frac{1}{2}, 1)) = 1$.

Assume now that $\alpha > 1$ in which case μ_α is infinite. In this case we use (7) together with a different estimate of R_n which also originates from [11, theorem 2.19]. In this case we have that for any $\kappa > 0$, the set of x for which the estimate

$$n^{\alpha-\kappa} \leqslant R_n(x) \leqslant n^{\alpha+\kappa} \tag{9}$$

does not hold is of arbitrarily small measure if n is sufficiently large. Let again N_0 be so large that the set D_δ of x for which (9) holds for $n > N_0$ is of measure at least $1-\delta$. For $x \in D_\delta$ we have that if there exists an $\varepsilon > 0$ such that $\tilde{\mu}_\alpha(B_m) \leqslant \frac{c}{m^{\frac{1}{\alpha}+\varepsilon}}$ for any $c > 0$, then

$$\tilde{\mu}_\alpha(B_{R_n}) \leqslant \frac{c}{R_n^{\frac{1}{\alpha}+\varepsilon}} \leqslant \frac{c}{nn^{\varepsilon\alpha-\frac{\kappa}{\alpha}-\varepsilon\kappa}} < \frac{c_1}{n} \quad , \quad c_1 < 1$$

for sufficiently large n when we choose κ sufficiently small compared to ε. As before we may pick $C_m := B(y, r_m)$ such that $\mu(C_m) = \frac{c_1}{m}$. As before we can define a set $\tilde{A_2}$ to which we can apply corollary 1 and by the same reasoning we conclude $\tilde{\mu}_\alpha(\mathcal{E}_{\text{ah}} \cap [\frac{1}{2}, 1)) = 0$.

The proof that $\tilde{\mu}_\alpha(B_m) \geqslant \frac{c}{m^{\frac{1}{\alpha}-\varepsilon}}$ implies $\tilde{\mu}_\alpha(\complement\mathcal{E}_{\text{ah}} \cap [\frac{1}{2}, 1)) = 0$ is similar to the previous two cases and we omit the details.

We now argue why our results for $\tilde{\mu}_\alpha(\mathcal{E}_{\text{ah}}(\mathbf{B}) \cap [\frac{1}{2}, 1))$ imply the general statement. Notice first that each assumption on $\mu(B_m)$ in Theorem 3(b), (c) and (d) is invariant under multiplication by a constant. This also yields for any $k \in \mathbb{N}$ that if a sequence of targets $\{B_m\}$ satisfies the assumption in Theorem 3(b), (c) or (d), then the sequence $\{B'_m\}$ with $B'_m = B_{m+k}$ satisfies that assumption as well.

For a start, we still consider the case of center $y \in [\frac{1}{2}, 1)$. Note that for every $x \in (0, \frac{1}{2})$ there is a smallest positive integer $k(x)$ such that $g_\alpha^{k(x)}(x) \in [\frac{1}{2}, 1)$. Then $x \in \mathcal{E}_{\text{ah}}(\{B_m\})$ if and only if $g_\alpha^{k(x)}(x) \in \mathcal{E}_{\text{ah}}(\{B_{m+k(x)}\}) \cap [\frac{1}{2}, 1)$. Hence, we have $\mathcal{E}_{\text{ah}}(\{B_m\}) \subset \bigcup_{k\in\mathbb{N}} g_\alpha^{-k}\left(\mathcal{E}_{\text{ah}}(\{B_{m+k}\}) \cap [\frac{1}{2}, 1]\right)$. Under the assumption from part (c) we conclude that $\mathcal{E}_{\text{ah}}(\{B_m\})$ is a null set since it is a countable union of null sets by our previous observations. The identical argument works for parts (b) and (d) when $\mathcal{E}_{\text{ah}}$ is replaced by $\complement\mathcal{E}_{\text{ah}}$.

To go from center $y \in [\frac{1}{2}, 1)$ to $y \in (0, 1)$ requires only a little more consideration. We again use that each assumption on $\mu(B_m)$ in theorem 3(b), (c) and (d) is invariant under multiplication by a constant. Assume for example the setting of theorem 3(c), i.e. $\alpha \geqslant 1$ and $\mu(B_m) \leqslant \frac{c}{m^{\frac{1}{\alpha}+\varepsilon}}$ for some $\varepsilon > 0$ and any $c > 0$. Assume that $y \in (0, \frac{1}{2})$. Let k_0 denote the smallest number such that $g_\alpha^{k_0}(y) \in [\frac{1}{2}, 1)$. There exists a $K > 0$ such that for all $m \in \mathbb{N}$ we have $\mu_\alpha(g_\alpha^{k_0}(B(y, r_m))) \leqslant K\mu_\alpha(B(y, r_m))$. This is an easy consequence of l'Hôpitals rule applied to the function $f(r) := \mu_\alpha(g_\alpha(B(y, r)))/\mu_\alpha(B(y, r))$. Pick $\tilde{B}_m$ to be the smallest ball with center $g_\alpha^{k_0}(y)$ such that $g_\alpha^{k_0}(B(y, r_m)) \subset \tilde{B}_m$. Then $\tilde{B}_m$ also satisfies the assumption of theorem (c) and we know that $\mu_\alpha(\mathcal{E}_{\text{ah}}(\tilde{B}_m)) = 0$ from the arguments above.

We argue that $\mathcal{E}_{\text{ah}}(B_m) \subset \mathcal{E}_{\text{ah}}(\tilde{B}_m)$. Assume that $x \in \mathcal{E}_{\text{ah}}(B_m)$, i.e. $\exists m_0 \forall m \geqslant m_0 \exists k < m : g_\alpha^k(x) \in B_m$. But if $g_\alpha^k(x) \in B_m$ then $g_\alpha^{k+k_0}(x) \in \tilde{B}_m$. Hence $\forall m \geqslant m_0 + k_0 \exists k < m$ such that $g_\alpha^k(x) \in \tilde{B}_m$, i.e. $x \in \mathcal{E}_{\text{ah}}(\tilde{B}_m)$. The cases theorem 3(b) and (d) follow by analogue arguments. This completes the proof of theorem 3.

4.3. The Gauß map

In this section, we will consider the Gauß map $G\colon (0, 1] \to [0, 1)$ defined by $G(x) = \frac{1}{x} \bmod 1$. This is a piecewise expanding map with infinitely many branches. There is a unique measure which is an invariant probability measure and absolutely continuous with respect to the Lebesgue measure. We denote this so-called Gauß measure by μ, and its density with respect to Lebesgue measure is given by $h(x) = \frac{1}{\log 2}\frac{1}{1+x}$.

HTS for the Gauß map is known for $B_m := [0, r_m)$, i.e. the interval with fixed left endpoint being 0 and right endpoint shrinking towards 0. More precisely, for $E_m = [0, r_m]$ we have,

[7] To be precise, Doeblins proof contained a gap which was repaired by Iosifescu [13] in 1977, but not before Galambos [10] had proven a special case 1972 which is sufficient for our purposes.

$$G_{E_m}(t) = 1 - \mathrm{e}^{-t}.$$

This follows from a classic result of Doeblin[7] [5] which may easily be translated into the above, see for example [12, section 5]. Hence for this choice of targets, corollary 2 applies. In particular, we have $\mu(\mathcal{E}_{\mathrm{ah}}) = 0$ provided $\mu([0, r_m)) \leqslant \frac{c_0}{m}$ for some constant $c_0 > 0$.

In order to show that we may apply theorem 5 we need the following short argument. Define $g : [0, 1] \to [0, 1]$ by $g(x) = \frac{1}{|G'(x)|} = x^2$ whenever x is not 0 or of the form $x = \frac{1}{n}$ for some $n \in \mathbb{N}$. In the remaining points, we let $g(x) = 0$. Obviously, g is then of bounded variation, since the heights of the jumps at $\frac{1}{n}$ are summable.

We have $||g||_\infty = 1$, but $||g \cdot (g \circ G)||_\infty < 1$, since $g(x) \to 1$ only if $x \to 1$, and $x = 1$ is not a fixed point under G. This implies that G^2 together with $g_2 = g \cdot g \circ G$ satisfies the assumptions of Rychlik [23]. Hence we may conclude that if $P \colon BV \to BV$ denotes the transfer operator associated to G, then, since G is mixing, there is only one eigenvalue on the unit circle, and P can be written in the form $P = Q + R$, where Q is the projection on the invariant density h, R has a spectral radius which is strictly less than 1 and $QR = RQ = 0$ [23, theorem 1 (c)]. That Q is the projection on the invariant density means that

$$Qf = h \int f \, \mathrm{d}x.$$

Let $f \in L^1$ and $g \in BV$. Let $c = -\int g \, \mathrm{d}\mu$, so that $\int (g + c) \, \mathrm{d}\mu = 0$. Then

$$\int f \circ G^n g \, \mathrm{d}\mu = \int f \circ G^n (g + c) \, \mathrm{d}\mu + \int f \, \mathrm{d}\mu \int g \, \mathrm{d}\mu.$$

By the choice of c, we have $Q((g + c)h) = 0$. Hence

$$\int f \circ G^n \cdot (g + c) \, \mathrm{d}\mu = \int f \cdot R^n((g + c)h) \, \mathrm{d}x.$$

Since R has a spectral radius strictly less than 1, there are positive constants C_1 and τ such that $||R^n((g + c)h)||_{BV} \leqslant C\mathrm{e}^{-\tau n}||(g + c)h||_{BV}$. In particular

$$\begin{aligned} ||R^n((g + c)h)||_\infty &\leqslant ||R^n((g + c)h)||_{BV} \leqslant C_1 \mathrm{e}^{-\tau n}||(g + c)h||_{BV} \\ &\leqslant C_1 \mathrm{e}^{-\tau n}(||gh||_{BV} + ||ch||_{BV}) \leqslant C_2 \mathrm{e}^{-\tau n}||g||_{BV}. \end{aligned}$$

From this, it follows that

$$\begin{aligned} \left| \int f \cdot R^n((g + c)h) \, \mathrm{d}x \right| &\leqslant \int |f| \, \mathrm{d}x \cdot ||R^n((g + c)h)||_\infty \\ &\leqslant \frac{||f||_1}{2 \log 2} \cdot C_2 \mathrm{e}^{-\tau n}||g||_{BV}. \end{aligned}$$

Hence, with $C = \frac{C_2}{2 \log 2}$ we have

$$\left| \int f \circ G^n g \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu \int g \, \mathrm{d}\mu \right| \leqslant C\mathrm{e}^{-\tau n}||f||_1 ||g||_{BV}.$$

In conclusion, we may apply theorem 5 to the Gauß map. In particular, we have $\mu(\mathcal{E}_{\mathrm{ah}}) = 1$ provided $\mu(B_m) \geqslant \frac{c_0 (\log m)^2}{m}$ for some constant $c_0 > \tau^{-1}$.

In order to obtain theorem 4 we now use the fact that when r_m is small and $G^{j-1}(x) \in [0, r_m)$ then $a_j(x) \sim \frac{1}{G^{j-1}(x)}$. An easy calculation shows that in the definition of $\mathcal{E}_{\mathrm{ah}}([0, r_m))$ for the Gauß map we can replace $\left\{G^k(x)\right\}_{k=0}^{m-1}(x) \cap [0, r_m) \neq \emptyset$ with

$\{a_k(x)\}_{k=1}^m \cap \left[\frac{1}{r_m}, \infty\right) \neq \emptyset$. The only thing left to do now is to compute the bounds on r_m when $\mu([0, r_m)) \leqslant \frac{c_0}{m}$ and $\mu([0, r_m)) \geqslant \frac{c_0(\log m)^2}{m}$. Using the density of the Gauß measure we get that

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

This leads to the following conclusions. If $r_m \leqslant \mathrm{e}^{\frac{c_0 \log 2}{m}} - 1$ for some $c_0 > 0$, then

$$\mu\Big(\Big\{\, x \in [0, 1] : \exists\, m_0 \in \mathbb{N}\, \forall m \geqslant m_0 : \max_{1 \leqslant k \leqslant m} a_k(x) \geqslant r_m^{-1} \,\Big\}\Big) = 0, \tag{10}$$

and if $r_m \geqslant \mathrm{e}^{\frac{c_0 \log 2(\log m)^2}{m}} - 1$ for some $c_0 > 0$ sufficiently large, then

$$\mu\Big(\Big\{\, x \in [0, 1] : \exists\, m_0 \in \mathbb{N}\, \forall m \geqslant m_0 : \max_{1 \leqslant k \leqslant m} a_k(x) \geqslant r_m^{-1} \,\Big\}\Big) = 1. \tag{11}$$

Since $\frac{c_0 \log 2}{m} \leqslant \mathrm{e}^{\frac{c_0 \log 2}{m}} - 1 \leqslant \frac{2c_0 \log 2}{m}$ if m is large and since $c_0 > 0$ arbitrary, we can conclude that (10) holds provided $r_m \leqslant \frac{1}{cm}$ for some $c > 0$. Letting $b_m = r_m^{-1}$, we have that

$$\mu\Big(\Big\{\, x \in [0, 1] : \exists\, m_0 \in \mathbb{N}\, \forall m \geqslant m_0 : \max_{1 \leqslant k \leqslant m} a_k(x) \geqslant b_m \,\Big\}\Big) = 0,$$

if $b_m \geqslant cm$ for some $c > 0$. This is statement theorem 4(a).

Similarly, we may conclude that (11) holds provided $r_m \geqslant \frac{(\log m)^2}{cm}$ for some sufficiently small $c > 0$. Letting $b_m = r_m^{-1}$ we then have

$$\mu\Big(\Big\{\, x \in [0, 1] : \exists\, m_0 \in \mathbb{N}\, \forall m \geqslant m_0 : \max_{1 \leqslant k \leqslant m} a_k(x) \geqslant b_m \,\Big\}\Big) = 1,$$

if $b_m \leqslant \frac{cm}{(\log m)^2}$ for some sufficiently small $c > 0$. This is theorem 4(b).

5. Results on the packing dimension

In this section we consider the case when $T = T_\beta$ is the β-transformation $[0, 1) \to [0, 1)$, defined for $\beta > 1$ by $T_\beta(x) = \beta x \bmod 1$, and μ is the unique invariant probability measure which is equivalent to Lebesgue measure. If we put $\Sigma = \{0, 1, \ldots, \beta - 1\}^{\mathbb{N}}$ in case of β an integer and $\Sigma = \{0, 1, \ldots, \lfloor\beta\rfloor\}^{\mathbb{N}}$ otherwise, then every $x \in [0, 1)$ can be coded by a sequence $(x_i) = (d_i(x))_{i=0}^{\infty}$ such that

$$d_i(x) = \begin{cases} 0 & \text{if } T^i(x) \in [0, \frac{1}{\beta}), \\ 1 & \text{if } T^i(x) \in [\frac{1}{\beta}, \frac{2}{\beta}), \\ \ldots & \ldots \\ \lfloor\beta\rfloor & \text{if } T^i(x) \in [\frac{\lfloor\beta\rfloor}{\beta}, 1). \end{cases}$$

Unless β is an integer, not every sequence of Σ occurs in this way. The closure (in the product topology) of $d([0, 1))$ is called the β-shift and is denoted by S_β.

Note that in the symbolic setting the doubling map becomes the left shift σ on Σ, i.e. for $d(x) = x_0, x_1, \ldots$ we have $d(T_\beta(x)) = \sigma(d(x)) = x_1, x_2, \ldots$. Given a finite sequence $x_0, x_1, \ldots, x_m$, we let

$$C(x_0, x_1, \ldots, x_m) = \{\, x \in [0, 1) : d_i(x) = x_i \text{ for } n = 0, 1, \ldots, m \,\}.$$

Then $C(x_0, x_1, \ldots, x_m)$ is an interval of length at most $\beta^{-(m+1)}$. (If β is an integer, then all intervals $C(x_0, x_1, \ldots, x_m)$ are of the same length $\beta^{-(m+1)}$. Otherwise these intervals are of different lengths.)

We recall the construction of the packing dimension. Let $F \subset \mathbb{R}^d$ and let $\delta > 0$. A collection $\{B_i\}$ of disjoint balls of radii at most δ with centres in F is called a *δ-packing collection* for F. For $s \geqslant 0$, let

$$\mathcal{P}^s_\delta(F) = \sup \left\{ \sum_i |B_i|^s : \{B_i\} \text{ is a } \delta\text{-packing collection for } F \right\}$$

and

$$\mathcal{P}^s_0(F) = \lim_{\delta \to 0} \mathcal{P}^s_\delta(F).$$

We then define the s-dimensional *packing measure* by

$$\mathcal{P}^s(F) = \inf \left\{ \sum_i \mathcal{P}^s_0(F_i) : F \subset \bigcup_i F_i \right\}$$

where the infimum is taken over all countable covers of F. Finally, the *packing dimension* of F is defined by

$$\dim_{\mathrm{P}} F = \sup \{ s : \mathcal{P}^s(F) = \infty \} = \inf \{ s : \mathcal{P}^s(F) = 0 \}.$$

In the hierarchy of dimensions the packing dimension falls between the Hausdorff dimension and the upper box-counting dimension in the sense that

$$\dim_{\mathrm{H}} F \leqslant \dim_{\mathrm{P}} F \leqslant \overline{\dim}_B F.$$

Theorem 6. *Suppose $B_m = B(0, \beta^{-sm})$, where $s \in (0, 1)$. Then,*

$$\dim_{\mathrm{P}} \mathcal{E}_{\mathrm{ah}} = 1 - s.$$

If $s \geqslant 1$ then $\dim_{\mathrm{P}} \mathcal{E}_{\mathrm{ah}} = 0$. *Indeed, $\mathcal{E}_{\mathrm{ah}}$ is a countable set in this case.*

Proof. We will first give the proof in the case that $\beta = 2$. This case is somewhat easier since in this case we have $S_\beta = \Sigma$. When β is an integer larger than 2, the proof only needs notational changes, but when β is not an integer, a little more care has to be taken. We will explain in the end of the proof which changes are needed to cover also the case when β is not an integer.

If $s = 1$, then we will prove that $\mathcal{E}_{\mathrm{ah}}$ consists only of those $x \in [0, 1)$, such that $T^n(x) = 0$ for some n. Hence $\mathcal{E}_{\mathrm{ah}}$ is the set of all finite words concatenated with an infinite tail of zeroes from which it is clear that $\mathcal{E}_{\mathrm{ah}}$ is countable. Then $\mathcal{E}_{\mathrm{ah}}$ must also be countable for all $s > 1$. Assume $x \in \mathcal{E}_{\mathrm{ah}}$. Then $x \in A_n$ for some n which means that $\left\{T^i x\right\}_{i=0}^{m-1} \cap B_m \neq \emptyset$ for all $m \geqslant n$. Then somewhere from digit number 0 to digit number $m - 1$, a block of m zeroes starts. Regardless of where this block starts it will overlap with the digit at place $m - 1$ and hence the digit on place $m - 1$ is 0. Since this is true for all $m \geqslant n$ we have shown that $d_m(x) = 0$ for all $m \geqslant n - 1$.

From now on, we assume that $s < 1$. We will first prove that $\dim_{\mathrm{P}} \mathcal{E}_{\mathrm{ah}} \leqslant 1 - s$. Since $\dim_{\mathrm{P}}$ is countably stable, $\mathcal{E}_{\mathrm{ah}} = \bigcup A_n$, $A_n \subset A_{n+1}$ and $\dim_{\mathrm{P}} \leqslant \dim_{\mathrm{B}}$, it is enough to prove that $\lim_{n\to\infty} \dim_{\mathrm{B}} A_n \leqslant 1 - s$, where $\dim_{\mathrm{B}}$ denotes the box dimension. If there is an n such that $T^n(x) = 0$, then $x \in A_n$. As in the introduction, let Λ denote the set of all such points which, as discussed above, is countable and may be disregarded, since the packing dimension of any countable set is 0. Set $A'_n := A_n \backslash \Lambda$.

We first prove that $x \in A'_n$ if and only if $(d_i(x))_{i=0}^{\infty}$ has blocks of $n_j + 1$ zeroes starting at position m_{j-1} for all $j \geqslant 1$ where $(n_j)_{j=0}^{\infty}$ and $(m_j)_{j=0}^{\infty}$ are strictly increasing sequences satisfying

(a) $m_0 < n$.
(b) $n_j \geqslant sm_j$.
(c) $m_{j-1} < (1-s)m_j$.

Suppose first that $x \in A'_n$. Then starting somewhere not later than at position $n-1$, the sequence $(d_i(x))_{i=0}^{\infty}$ contains a block of at least sn zeroes[8]. We let n_1 be the length of the largest such block, and let $m_0 < n$ be the position of its first zero. This satisfies (a).

The above property is true for all $m \geqslant n$, i.e. starting somewhere not later than at position $m-1$ a block of at least sm zeroes start. This allows us to define sequences $(n_j)_{j=1}^{\infty}$ and $(m_j)_{j=0}^{\infty}$ as follows. The numbers n_1 and m_0 are already defined. Suppose that n_j and m_{j-1} are defined. Then we let m_j be the position of the first digit of the leftmost block of at least $n_j + 1$ consecutive zeroes in the sequence $(d_i(x))_{i=0}^{\infty}$ and we let n_{j+1} be the maximal number of zeroes in such a block. Note that, if we make any change in the digits $d_i(x)$ outside of the blocks of n_j zeroes described above, then we get a new sequence $d_i(y)$ with $y \in A'_n$.

Since $x \in A'_n$ and n_j was chosen to be maximal, the block of length n_j must be long enough to ensure that $\left\{T^i x\right\}_{i=0}^{m_j-1} \cap B_{m_j} \neq \emptyset$, i.e. we always have $n_j \geqslant sm_j$ and (b) is satisfied. Furthermore, again due to n_j being maximal and since $x \notin \Lambda$, the blocks of zeroes are separated, and we have $m_{j-1} + n_j < m_j$. Hence $m_{j-1} < (1-s)m_j$ and (c) is satisfied.

Conversely it is clear that any $x \in [0,1]$ for which (a)–(c) holds true for $(d_i(x))_{i=0}^{\infty}$ is an element of A'_n.

Let now N be fixed and take k such that $m_{k-1} < N \leqslant m_k$. From $m_{j-1} < (1-s)m_j$ we obtain that

$$1 < m_1 < (1-s)^{k-2} m_{k-1} < (1-s)^{k-2} N.$$

Hence

$$k < 2 + \frac{\log N}{-\log(1-s)}.$$

The numbers of zeroes in $(d_i(x))_{i=0}^{\infty}$ between digit number m_0 and digit number N is at least $s(N - m_0)$. This is the case, since the finite sequence $d_{m_0}(x), \ldots, d_N(x)$ can be cut into k sequences starting at $d_{m_j}(x)$, $j = 0, 1, \ldots k-1$ and on each of these subsequences a proportion of at least s of the length consists of zeroes.

We will now cover the set A'_n by intervals of the form

$$C(x_0, x_1, \ldots, x_N).$$

The sequences $x_0, x_1, \ldots, x_N$ that we need to consider are only those that can be obtained from a sequence $(n_j)_{j=1}^{k}$ and $(m_j)_{j=0}^{k}$ satisfying the inequalities (a)–(c) and therefore also with $k \leqslant 2 + \frac{\log N}{-\log(1-s)}$.

The sequence $(m_j)_{j=0}^{k}$ can be chosen in at most N^{k+1} different ways, and once $(m_j)_{j=0}^{k}$ is chosen, we can choose $(n_j)_{j=1}^{k}$ in at most N^k different way, hence in total at most N^{2k+1} different ways to choose the sequences. (These are very rough estimates, but sufficient for our purpose.)

[8] Since sn is typically non-integer we should, in principle, be more diligent and write $\lceil sn \rceil$. However, to improve readability we let the relevant ceiling and floor functions be implicitly understood throughout the proof. The outcome is invariant under this abuse of notation.

Once the sequences $(n_j)_{j=1}^k$ and $(m_j)_{j=0}^k$ are chosen, we have specified a certain number of zeroes, while the other digits in the sequence $x_0, x_1, \ldots, x_N$ remain free. There are at most $m_0 + (1-s)(N-m_0) \leqslant (1-s)N + n$ digits that are free, and hence once the sequences $(n_j)_{j=1}^k$ and $(m_j)_{j=0}^k$ are chosen, we may choose the sequence $x_0, x_1, \ldots, x_N$ in at most $2^{(1-s)N+n}$ ways.

In total, the number of sequences $x_0, x_1, \ldots, x_N$ that we need in order to cover A'_n with the sets $C(x_0, x_1, \ldots, x_N)$, are not more than

$$N^{2k+1} 2^{(1-s)N+n} \leqslant N^{5+2\frac{\log N}{-\log(1-s)}} 2^{(1-s)N+n}. \tag{12}$$

Since the sets $C(x_0, x_1, \ldots, x_N)$ have diameter $2^{-(N+1)}$ we get from the definition of box dimension that

$$\dim_{\mathrm{B}}(A'_n) \leqslant \lim_{N\to\infty} \frac{\log\left(N^{5+2\frac{\log N}{-\log(1-s)}} 2^{(1-s)N+n}\right)}{-\log\left(2^{-(N+1)}\right)} = 1 - s. \tag{13}$$

The last equality follows since $N^{\log N}$ grows with N slower than any exponential. Since the bound is independent of n it also holds as $n \to \infty$.

We now finish by proving that $\dim_{\mathrm{P}} \mathcal{E}_{\mathrm{ah}} \geqslant 1 - s$. Let $(m_j)_{j=1}^\infty$ be a strictly increasing sequence of natural numbers. Using this sequence, we will construct a subset of $\mathcal{E}_{\mathrm{ah}}$ and prove that the packing dimension of this subset is $1 - s$ if the sequence $(m_j)_{j=1}^\infty$ is chosen such that

$$\lim_{k\to\infty} \frac{\sum_{j=1}^k m_j}{m_k} = 1. \tag{14}$$

For example, we could choose the sequence inductively as $m_k := k \sum_{j=1}^{k-1} m_j$, however, the explicit choice is irrelevant. Let F consist of those $x \in [0, 1)$ such that

$$d_{m_j}(x), d_{m_j+1}(x), \ldots, d_{m_j+sm_{j+1}}(x) = 0, 0, \ldots, 0$$

for all $j \geqslant 0$. Then $F \subset \mathcal{E}_{\mathrm{ah}}$.

We let μ be the probability measure on F defined by

$$\mu(C(x_0, x_1, \ldots, x_n)) = \frac{1}{N(n)},$$

if $C(x_0, x_1, \ldots, x_n)$ intersects F where $N(n)$ is the number of intervals $C(x_0, x_1, \ldots, x_n)$ that have non-empty intersection with F. Otherwise we assign measure 0. The upper pointwise dimension of μ at $x \in F$ is defined by

$$\overline{d}_\mu(x) = \limsup_{r\to 0} \frac{\log \mu\,(B(x, r))}{\log r}.$$

Let $r_n = 2^{-n-2}$. Then $B(x,r_n)$ is contained in the cylinder $C(x_0, x_1, \ldots, x_n)$ and one of the neigbouring cylinders. The measure of the neighbouring cylinder is either zero or equal to that of the cylinder $C(x_0, x_1, \ldots, x_n)$. Hence

$$\mu(B(x, r_n)) \leqslant 2\mu(C(x_0, x_1, \ldots, x_n)),$$

and since $\log r_n < 0$ we have

$$\frac{\mu(B(x, r_n))}{\log r_n} \geqslant \frac{\log(2\mu(C(x_0, x_1, \ldots, x_n)))}{\log r_n}.$$

It therefore follows that

$$\begin{aligned}\overline{d}_\mu(x) &= \limsup_{r\to 0} \frac{\log \mu\big(B(x,r)\big)}{\log r}\\ &\geqslant \limsup_{n\to\infty} \frac{\log \mu\big(B(x,r_n)\big)}{\log r_n}\\ &\geqslant \limsup_{n\to\infty} \frac{\log \mu\big(C(x_0,x_1,\ldots,x_n)\big)}{\log 2^{-(n+2)}}\\ &\geqslant \limsup_{k\to\infty} \frac{\log \mu\big(C(x_0,x_1,\ldots,x_{m_k})\big)}{\log 2^{-(m_k+2)}}.\end{aligned}$$

We will use this inequality to show that $\overline{d}_\mu(x) \geqslant 1-s$ for any $x \in F$. By [6, proposition 10.1], this implies that $\dim_{\mathrm{P}} F \geqslant 1-s$.

We have that

$$\mu(C(d_0(x), d_1(x), \ldots, d_{m_k}(x))) = \frac{1}{N(m_k)}, \tag{15}$$

and

$$N(m_k) = 2^{m_k - s\sum_{j=1}^{k} m_j}. \tag{16}$$

Hence

$$\overline{d}_\mu(x) \geqslant \lim_{k\to\infty} \frac{m_k - s\sum_{j=1}^{k} m_j}{m_k+2} = \lim_{k\to\infty} 1 - s\frac{\sum_{j=1}^{k} m_j}{m_k+2}.$$

Since m_k are chosen so that (14) holds, we have

$$\overline{d}_\mu(x) = 1-s,$$

which implies that $\dim_{\mathrm{P}} \mathcal{E}_{\mathrm{ah}} \geqslant \dim_{\mathrm{P}} F \geqslant 1-s$.

We will now comment on the changes needed when β is not an integer. In the proof of the upper bound, when constructing the cover, we need to make the following change. In the blocks between the blocks of zeroes, we only consider sequences which occur in S_β (rather than Σ). It is well known that the number of sequences of length n in S_β is approximately β^n, and certainly less than $(\beta+\varepsilon_n)^n$, for some sequence $\varepsilon_n \to 0$ as $n\to\infty$. This leads to the estimate that we can cover A'_n by

$$N^{5+2\frac{\log N}{-\log(1-s)}} (\beta+\varepsilon_n)^{(1-s)N+n}$$

sets of the form $C(x_0, x_1, \ldots, x_N)$ (compare with (12)). The sets used in the cover do not have to be of the same size, but making some larger so that they are all of diameter $\beta^{-(N+1)}$ we still have a cover which implies that the box dimension can be estimated by

$$\dim_{\mathrm{B}}(A'_n) \leqslant \lim_{N\to\infty} \frac{\log\big(N^{5+2\frac{\log N}{-\log(1-s)}} (\beta+\varepsilon_n)^{(1-s)N+n}\big)}{-\log\big(\beta^{-(N+1)}\big)} = (1-s)\frac{\log(\beta+\varepsilon_n)}{\log\beta}.$$

(Compare with (13).) Letting $n\to\infty$ gives us $\lim_{n\to\infty} \dim_{\mathrm{B}}(A'_n) \leqslant 1-s$.

To obtain the lower bound, let F be the set of those $x \in [0,1)$ such that $(d_i(x))_{i=0}^{\infty} \in S_\beta$ and

$$d_{m_j}(x), d_{m_j+1}(x), \ldots, d_{m_j+sm_{j+1}}(x) = 0, 0, \ldots, 0$$

for all $j \geqslant 0$. We again construct a measure as in (15). However, computing $N(m_k)$ exactly is not as simple for general β as in (16). But we may bound this number from below and thereby get an upper bound on $\overline{d}_\mu(x)$ at $x \in F$. Analogue to the upper bound we may estimate

$$N(m_k) \geqslant (\beta - \varepsilon_k)^{m_k - s\sum_{j=1}^{k} m_j}$$

for some sequence $\varepsilon_k \to 0$ for $k \to \infty$.

Finally, consider a ball $B(x,r_n)$, where $r_n = \beta^{-n-2}$. The ball $B(x,r_n)$ is then contained in $2n$ consequtive cylinders $C(x_0, x_1, \ldots, x_n)$ [22, lemma 3]. Hence

$$\mu(B(x,r_n)) \leqslant 2n\mu(C(x_0, x_1, \ldots, x_n))$$

and the same computations as in the proof leads to

$$\overline{d}_\mu(x) \geqslant (1 - s).$$

Acknowledgments

We thank both Jörg Schmeling and Dmitry Kleinbock for interesting discussions and helpful comments and suggestions which improved the paper. We thank Jorge Freitas for answering our questions about Hitting Time Statistics for the Gauß map. We also acknowledge financial support by the Hamburg–Lund Funding Program 2018 which made several mutual research visits possible. Finally we thank Gerardo González Robert and the referees for their careful reading of the paper and helpful comments and corrections which improved the paper substantially.

References

[1] Aspenberg M and Persson T 2019 Shrinking targets in parametrised families *Math. Proc. Camb. Phil. Soc.* **166** 265–95
[2] Athreya J S 2009 Logarithm laws and shrinking target properties *Proc. Math. Sci.* 119 541–57
[3] Bugeaud Y and Liao L 2016 Uniform diophantine approximation related to b-ary and β-expansions *Ergod. Theor. Dynam. Syst.* **36** 1–22
[4] Chernov N and Kleinbock D 2001 Dynamical Borel–Cantelli lemmas for Gibbs measures *Isr. J. Math.* **122** 1–27
[5] Doeblin W 1940 Remarques sur la théorie métrique des fractions continues *Compos. Math.* **7** 353–71
[6] Falconer K 1997 *Techniques in Fractal Geometry* (New York: Wiley)
[7] Freitas A, Freitas J and Todd M 2010 Hitting time statistics and extreme value theory *Probab. Theor. Relat. Fields* **147** 675–710
[8] Freitas A, Freitas J and Todd M 2012 Extremal index, hitting time statistics and periodicity *Adv. Math.* **231** 2626–65
[9] Freitas A, Freitas J, Todd M and Vaienti S 2016 Rare events for the Manneville–Pomeau map *Stoch. Process. Appl.* **126** 3463–79
[10] Galambos J 1972 The distribution of the largest coefficient in continued fraction expansions *Q. J. Math.* **23** 147–51
[11] Galatolo S, Holland M, Persson T and Zhang Y 2018 Birkhoff sums of infinite observables and anomalous time-scaling of extreme events in infinite systems (arXiv:1810.10742)
[12] Haydn N 2013 Entry and return times distribution *Dynam. Syst.* **28** 333–53
[13] Iosifescu M 1977 A Poisson law for ψ-mixing sequences establishing the truth of a Doeblin's statement *Rev. Roum. Math. Pures Appl.* **22** 1441–7

[14] Kelmer D 2017 Shrinking targets for discrete time flows on hyperbolic manifolds *Geom. Funct. Anal.* **27** 1257–87
[15] Kelmer D and Yu S 2019 Shrinking targets problems for flows on homogeneous spaces *Trans. Amer. Math. Soc.* **372** 6283–314
[16] Kleinbock D, Konstantoulas I and Richter F K 2019 Zero–one laws for eventually always hitting points in mixing systems (arXiv:1904.08584)
[17] Kleinbock D and Wadleigh N 2019 An inhomogeneous Dirichlet theorem via shrinking targets *Compositio Mathematica* **155** 1402–23
[18] Liverani C 1995 Decay of correlations for piecewise expanding maps *J. Stat. Phys.* **78** 1111–29
[19] Liverani C, Saussol B and Vaienti S 1998 Conformal measure and decay of correlation for covering weighted systems *Ergod. Theor. Dynam. Syst.* **18** 1399–420
[20] Lucarini V, Faranda D, Freitas A, Freitas J, Holland M, Kuna T, Nicol M, Todd M and Vaienti S 2016 *Extremes and Recurrence in Dynamical Systems* (*Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts*) (Hoboken, NJ: Wiley)
[21] Oh H and Kelmer D 2018 Exponential mixing and shrinking targets for geodesic flow on geometrically finite hyperbolic manifolds (arXiv:1812.05251)
[22] Persson T and Schmeling J 2008 Dyadic diophantine approximation and Katok's horseshoe approximation *Acta Arith.* **132** 205–30
[23] Rychlik M 1983 Bounded variation and invariant measures *Stud. Math.* **76** 69–80
[24] Thaler M 1980 Estimates of the invariant densities of endomorphisms with indifferent fixed points *Isr. J. Math.* **37** 303–14
[25] Tseng J 2008 On circle rotations and the shrinking target properties *Discrete Contin. Dynam. Syst.* **20** 1111–22
[26] Young L-S 1992 Decay of correlations for certain quadratic maps *Commun. Math. Phys.* **146** 123–38