Ergod. Th. & Dynam. Sys. (2012), **32**, 1444–1470 © *Cambridge University Press, 2011*
doi:10.1017/S0143385711000368

The packing spectrum for Birkhoff averages on a self-affine repeller

HENRY W. J. REEVE

Department of Mathematics, The University of Bristol, University Walk, Clifton, Bristol BS8 1TW, UK
(e-mail: henrywjreeve@googlemail.com)

(*Received* 22 *June* 2010 *and accepted in revised form* 15 *March* 2011)

Abstract. We consider the multifractal analysis of Birkhoff averages of continuous potentials on a self-affine Sierpiński sponge. In particular, we give a variational principle for the packing dimension of the level sets. Furthermore, we prove that the packing spectrum is concave and continuous. We give a sufficient condition for the packing spectrum to be real analytic, but show that for general Hölder continuous potentials, this need not be the case. We also give a precise criterion for when the packing spectrum attains the full packing dimension of the repeller. Again, we present an example showing that this is not always the case.

1. *Introduction and statement of results*

Let Λ be the repeller of a $C^{1+\epsilon}$ map $f : X \to X$. Given some continuous potential $\varphi : \Lambda \to \mathbb{R}^N$ and some $\alpha \in \mathbb{R}^N$, we are interested in the set of points in the repeller for which the Birkhoff average converges to α:

$$E_\varphi(\alpha) := \left\{ x \in \Lambda : \lim_{n\to\infty} \frac{1}{n} \sum_{q=0}^{n-1} \varphi(f^q(x)) = \alpha \right\}. \tag{1.1}$$

We would like to understand how the geometric complexity of $E_\varphi(\alpha)$ varies as a function of α. Geometric complexity, here, is to be understood in terms of the dual notions of Hausdorff dimension $\dim_{\mathcal{H}}$, defined in terms of minimal coverings, and packing dimension $\dim_{\mathcal{P}}$, defined in terms of maximal packings (see [**Ed1**, §6.2] or [**Fa2**, Ch. 3]). We refer to $\alpha \mapsto \dim_{\mathcal{H}} E_\varphi(\alpha)$ as the Hausdorff spectrum and $\alpha \mapsto \dim_{\mathcal{P}} E_\varphi(\alpha)$ as the packing spectrum.

In the conformal setting there is a well-known variational principle giving the values for both spectra. To recall this result, we require some terminology. Let Λ be a repeller for an expanding $C^{1+\epsilon}$ map f of a smooth manifold X. We let $\mathcal{M}(\Lambda, f)$ denote the set of f-invariant Borel probability measures supported on Λ. Given $\mu \in \mathcal{M}(\Lambda, f)$, we let $h_\mu(f)$ denote the Kolmogorov–Sinai entropy (see [**Wa**, §4.10]) and let $\lambda_\mu(f) := \int \log \|f'\| \, d\mu$ denote the Lyapunov exponent. Given a continuous potential $\varphi : \Lambda \to \mathbb{R}$,

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

we let $A(\varphi) := \{\int \varphi \, d\mu : \mu \in \mathcal{M}(\Lambda, f)\}$. We may now recall the classic result of Pesin and Weiss [**PW1**], Fan *et al* [**FFW**], Barreira and Saussol [**BSa**], Feng *et al* [**FLW**] and Olsen [**Ol1, Ol3**]. See also [**GR, JJOP**] for extensions to classes of non-uniformly hyperbolic functions.

THEOREM 1. (Feng, Lau, Wu) *Let Λ be the repeller for an expanding $C^{1+\epsilon}$ map $f : X \to X$. Suppose that f is conformal and topologically mixing on Λ. Then, for all $\alpha \in A(\varphi)$, we have*

$$\dim_{\mathcal{H}} E_\varphi(\alpha) = \dim_{\mathcal{P}} E_\varphi(\alpha) = \sup\left\{\frac{h_\mu(f)}{\lambda_\mu(f)} : \mu \in \mathcal{M}(\Lambda, f), \int \varphi \, d\mu = \alpha\right\}.$$

In particular, when f is conformal and uniformly hyperbolic, the Hausdorff and packing spectra coincide. This is a consequence of the fact that the level set $E_\varphi(\alpha)$ corresponds to a type of statistical convergence, together with the neat relationship between geometric and statistical properties which holds in the conformal setting. By contrast, the packing and Hausdorff dimensions of level sets defined in terms of divergent asymptotic properties may differ (see [**BOS, Ol3**]). Theorem 1 also allows one to deduce various regularity properties of the spectrum [**BSa, FFW, Ol1, PW2**]. The spectrum is continuous and when φ is Hölder continuous, it is also real analytic. When the Lyapunov exponent $\lambda_\mu(f)$ is given by a fixed constant, independent of $\mu \in \mathcal{M}(\Lambda, f)$, the spectrum is also concave.

The dimension theory of non-conformal systems, for which there is no simple correspondence between geometric and statistical properties, is much less well understood. For the most part we only have almost all type results, both for the dimension of a repeller [**Fa1**] and for the dimension of level sets for Birkhoff averages [**JS**]. One class of non-conformal fractals for which we do have deterministic results [**BM1, Be, JR, Ki, KP, McM, Ni**] is the class of self-affine Sierpiński sponges introduced by Bedford [**Be**] and McMullen [**McM**], and generalized by Kenyon and Peres [**KP**].

DEFINITION 1.1. (Self-affine Sierpiński sponges) Let $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$ denote the d-dimensional torus. Choose natural numbers $a_1 > a_2 > \cdots > a_d \geq 2$. Let f denote the integer-valued diagonal map given by

$$(x_q)_{q=1}^d \mapsto (a_q x_q)_{q=1}^d \quad \text{for } (x_q)_{q=1}^d \in \mathbb{T}^d. \tag{1.2}$$

Given a digit set $\mathcal{D} \subseteq \prod_{q=1}^d \{0, \ldots, a_q - 1\}$, there is a corresponding self-affine repeller Λ given by

$$\Lambda := \left\{\left(\sum_{\nu=1}^{\infty} \frac{i_q(\nu)}{a_q^\nu}\right)_{q=1}^d : (i_q(\nu))_{q=1}^d \in \mathcal{D} \text{ for all } \nu \in \mathbb{N}\right\}. \tag{1.3}$$

A limit set Λ defined in this way is referred to as a self-affine Sierpiński sponge. A two-dimensional Sierpiński sponge is known as a Bedford–McMullen carpet (see Figure 1).

To state the relevant results concerning self-affine Sierpiński sponges we first introduce some terminology. Given a continuous transformation T of a metric space X, we let $\mathcal{M}(X, T)$ denote the set of T-invariant Borel probability measures, and let $\mathcal{E}(X, T)$ denote the set of $\mu \in \mathcal{M}(X, T)$ which are ergodic.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

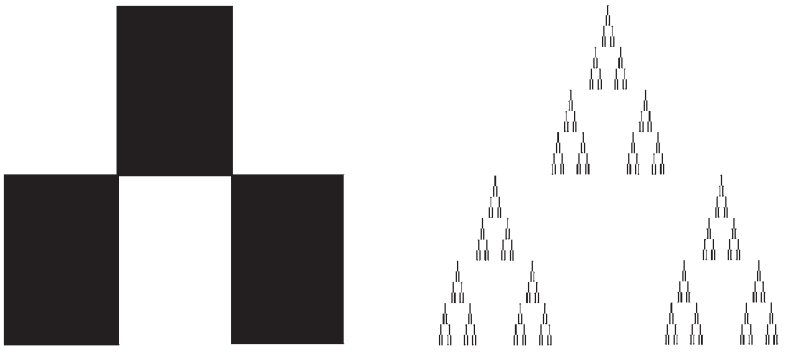


FIGURE 1. A representation of a digit set (left) and the corresponding Bedford–McMullen carpet (right).

Given $k \leq d$, we let $\pi_k : \mathbb{T}^d \to \mathbb{T}^{d-(k-1)}$ be the projection $\pi_k : (x_q)_{q=1}^d \mapsto (x_q)_{q=k}^d$. We let $f_k : \mathbb{T}^{d-(q-1)} \to \mathbb{T}^{d-(q-1)}$ denote the map $(x_q)_{q=k}^d \mapsto (a_q x_q)_{q=k}^d$.

Suppose $A \subseteq \mathcal{M}(\Lambda, f)$. For each $k \leq d$ we define $H^k(f, A)$ by

$$H^k(f, A) := \sup\{h_{\mu\circ\pi_k^{-1}}(f_k) : \mu \in A\}.$$

When $A = \mathcal{M}(\Lambda, f)$, we write $H^k(f, A) = H^k(f)$.

The following result is due to Kenyon and Peres [**KP**].

THEOREM 2. (Kenyon, Peres) *Let Λ be a self-affine Sierpiński sponge. Then,*

$$\dim_{\mathcal{H}}(\Lambda) = \sup_{\mu\in\mathcal{M}(\mathbb{T}^d, f)} \left\{ \frac{h_\mu(f)}{\log a_1} + \sum_{k=2}^{d} \left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}} \right) h_{\mu\circ\pi_k^{-1}}(f_k) \right\},$$

$$\dim_{\mathcal{P}}(\Lambda) = \frac{H^1(f)}{\log a_1} + \sum_{k=2}^{d} \left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}} \right) H^k(f).$$

Bedford [**Be**] and McMullen [**McM**] independently determined both the Hausdorff dimension and the upper box dimension in the two-dimensional setting. In [**KP**], Kenyon and Peres extend these results to higher dimensions. It follows from [**Fa3**, Proposition 3.6] that the formula for the upper box dimension also gives an expression for the packing dimension.

The multifractal analysis of Birkhoff averages is closely related to the multifractal analysis of pointiwise dimension. Given an invariant measure $\nu \in \mathcal{M}(\Lambda, f)$ on a self-affine Sierpiński sponge Λ, the Hausdorff and packing spectra for pointwise dimension are given by $\alpha \mapsto \dim_{\mathcal{H}} D_\nu(\alpha)$ and $\alpha \mapsto \dim_{\mathcal{P}} D_\nu(\alpha)$, respectively, where

$$D_\nu(\alpha) := \left\{ x \in \Lambda : \lim_{r\to 0} \frac{\log \nu(B(x, r))}{\log r} = \alpha \right\}. \tag{1.4}$$

In [**Ki**], King determined the Hausdorff spectrum for Bernoulli measures on a Bedford–McMullen carpet with strong separation conditions. Olsen extended King's result to Bernoulli measures on d-dimensional self-affine Sierpiński sponges [**Ol4**]. The Hausdorff spectrum for Gibbs measures was determined by Barral and Mensi [**BM2**] for Bedford–McMullen carpets, and by Barral and Feng [**BF**] for a d-dimensional self-affine

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Sierpiński sponge. In [**JR**], Jordan and Rams gave the Hausdorff spectrum for Bernoulli measures on a Bedford–McMullen carpet without the strong separation conditions required in [**BF, BM2, Ki, Ol4**]. In contrast, almost nothing is known about the packing spectrum for pointwise dimension on a self-affine Sierpiński sponge. In this paper we determine $\alpha \mapsto \dim_{\mathcal{P}} D_\nu(\alpha)$ for a very limited class of Bernoulli measures ν on self-affine Sierpiński sponges with strong separation conditions. We also give an example disproving a conjecture of Olsen [**Ol4**, Conjecture 4.1.7] (see §7). However, the main focus for this article is the packing spectrum for Birkhoff averages.

The first result concerning the multifractal analysis of Birkhoff averages for self-affine Sierpiński sponges is due to Nielsen [**Ni**]. Suppose Λ is a self-affine Sierpiński sponge. For $x \in \Lambda$, we let

$$\Gamma(x) := \left\{ ((i_q(\nu))_{q=1}^d)_{\nu\in\mathbb{N}} : x = \left(\sum_{\nu=1}^{\infty} \frac{i_q(\nu)}{a_q^\nu} \right)_{q=1}^d \right\}.$$

Given a probability vector $\mathbf{p} = (p_l)_{l\in\mathcal{D}}$, defined over a digit set $\mathcal{D}$, we let $N_l(\omega|n) := \#\{\nu \le n : \omega_\nu = l\}$, where # denotes cardinality, and define

$$\Lambda_{\mathbf{p}} := \left\{ x \in \Lambda : \exists \omega \in \Gamma(x) \text{ with } \lim_{n\to\infty} \frac{N_l(\omega|n)}{n} = p_l \text{ for each } l \in \mathcal{D} \right\}.$$

Let $\mu_{\mathbf{p}}$ denote the Bernoulli measure on Λ corresponding to the probability vector $\mathbf{p}$. In [**Ni**], Nielsen proved the following formula for the Hausdorff and packing dimensions of $\Lambda_{\mathbf{p}}$ in the two-dimensional case. With minor modifications, the proof also applies in higher dimensions.

THEOREM 3. (Nielsen) *Let Λ be a self-affine Sierpiński sponge with digit set $\mathcal{D}$ and let $\mathbf{p}$ be a probability vector defined over $\mathcal{D}$. Then,*

$$\dim_{\mathcal{P}} \Lambda_{\mathbf{p}} = \dim_{\mathcal{H}} \Lambda_{\mathbf{p}} = \frac{h_{\mu_{\mathbf{p}}}(f)}{\log a_1} + \sum_{k=2}^{d} \left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}} \right) h_{\mu_{\mathbf{p}}\circ\pi_k}(f_k).$$

In particular, this shows that for a certain special class of Birkhoff averages, defined over a self-affine Sierpiński sponge, we always have $\dim_{\mathcal{P}} E_\varphi(\alpha) = \dim_{\mathcal{H}} E_\varphi(\alpha)$. However, it follows from Theorem 2 that for self-affine Sierpiński sponges we often have $\dim_{\mathcal{H}}(\Lambda) < \dim_{\mathcal{P}}(\Lambda)$. Consequently, by considering any φ which is cohomologous to a constant, the Hausdorff and packing spectra for Birkhoff averages on a Bedford–McMullen repeller do not always coincide. This is a consequence of there being two distinct rates of expansion. It takes less time for a difference along the direction of strong repulsion to be blown up to the scale of the Markov partition than it does for a similarly sized difference along the direction of weak repulsion. As such a given geometric scale will correspond to two timescales, often resulting in a difference between Hausdorff and packing dimensions. The reason that this does not affect the coincidence of $\dim_{\mathcal{P}} \Lambda_{\mathbf{p}}$ and $\dim_{\mathcal{H}} \Lambda_{\mathbf{p}}$ is that the convergence of $N_d(\omega|n)/n$ forces points in the set $\Lambda_{\mathbf{p}}$ to display similar behaviour at both time scales. For less restrictive level sets this need not be the case.

This extra level of complexity in the non-conformal case makes the question of Hausdorff and packing spectra for Birkhoff averages on Bedford–McMullen repellers

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

an interesting one, where we do not expect to observe the same behaviour as in the conformal case. The first part of this question was answered by Barral, Feng and Mensi in [**BF, BM1**]. Given an integer-valued diagonal map f on a self-affine Sierpiński sponge Λ and a continuous potential $\varphi : \Lambda \to \mathbb{R}^N$, we let $A(\varphi) := \{\int \varphi \, d\mu : \mu \in \mathcal{M}(\Lambda, f)\}$. One can easily see that $E_\varphi(\alpha) = \emptyset$ for $\alpha \notin A(\varphi)$. The following result concerning the Hausdorff spectrum is due to Barral and Feng [**BF**].

THEOREM 4. (Barral, Feng) *Let Λ be a self-affine Sierpiński sponge. Let $\varphi : \Lambda \to \mathbb{R}^N$ be a continuous potential. Then, for all $\alpha \in A(\varphi)$, we have*

$$\dim_{\mathcal{H}} E_\varphi(\alpha) = \sup\left\{\frac{h_\mu(f)}{\log a_1} + \sum_{k=2}^{d}\left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}}\right) h_{\mu\circ\pi_k^{-1}}(f_k)\right\},$$

where the supremum is taken over all $\mu \in \mathcal{M}(\Lambda, f)$ with $\int \varphi \, d\mu = \alpha$.

This extends the work of Barral and Mensi in [**BM1**], where the Hausdorff spectrum for Hölder continuous potentials on a Bedford–McMullen carpet are given as the Legendre transform of an explicit moment function.

In this paper, we prove a dual result for the packing spectrum. For each $\alpha \in A(\varphi)$, we define $H^k(T, \varphi, \alpha)$ for $k = 1, \ldots, d$ by

$$H^k(f, \varphi, \alpha) := \sup\left\{h_{\mu\circ\pi_k^{-1}}(f_k) : \mu \in \mathcal{M}_T(\Lambda), \int \varphi \, d\mu = \alpha\right\}.$$

THEOREM 5. *Let Λ be a self-affine Sierpiński sponge. Let $\varphi : \Lambda \to \mathbb{R}^N$ be some continuous potential. Then, for all $\alpha \in A(\varphi)$, we have*

$$\dim_{\mathcal{P}} E_\varphi(\alpha) = \frac{H^1(f, \varphi, \alpha)}{\log a_1} + \sum_{k=2}^{d}\left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}}\right) H^k(f, \varphi, \alpha).$$

In fact, Theorem 5 follows from the more general Theorem 6. Given a Borel probability measure $\mu \in \mathcal{M}(\Lambda)$, we define

$$A_n(\mu) := \frac{1}{n}\sum_{k=0}^{n-1} \mu \circ f^{-k}. \tag{1.5}$$

Given $x \in \Lambda$, we let $\mathcal{V}(x)$ denote the set of all weak $*$ accumulation points of the sequence of measures $(A_n(\delta_x))_{n\in\mathbb{N}}$, where δ_x denotes the Dirac measure concentrated at x. Note that $\mathcal{V}(x) \subseteq \mathcal{M}(\Lambda, f)$ [**Wa**, Theorem 6.9] for all $x \in \Lambda$. Given $A \subseteq \mathcal{M}(\Lambda, f)$ we define

$$\begin{aligned} X(A) &:= \{x \in \Lambda : \mathcal{V}(x) = A\}, \\ Y(A) &:= \{x \in \Lambda : \mathcal{V}(x) \subseteq A\}. \end{aligned} \tag{1.6}$$

In [**BF**], Barral and Feng considered the special case in which $A = \{\mu\}$ for some $\mu \in \mathcal{M}(\Lambda, f)$. It follows that $X(\{\mu\}) = Y(\{\mu\})$ and the Hausdorff and packing dimensions coincide. However, in general, this is not the case.

THEOREM 6. *Let Λ be a self-affine Sierpiński sponge. Suppose that A is a non-empty closed convex subset of $\mathcal{M}(\Lambda, f)$. Then,*

$$\dim_{\mathcal{P}} X(A) = \dim_{\mathcal{P}} Y(A) = \frac{H^1(f, A)}{\log a_1} + \sum_{k=2}^{d}\left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}}\right) H^k(f, A).$$

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

The central difficulty in determining the packing spectrum is proving the lower bound in Theorem 6. Unlike the Hausdorff dimension, the packing dimension of a level set typically exceeds the supremum of the dimensions of the invariant measures supported on that set. We construct a non-invariant measure specifically suited to obtaining an optimal lower bound for packing dimension.

The rest of the paper is structured as follows. We begin by restating Theorems 5 and 6 in §2 in terms of the symbolic space. The proof of Theorem 7 is given in §§3 and 4. In §3, we prove the lower bound, and in §4, we prove the upper bound. In §5, we deduce some regularity properties of the packing spectrum. In §6, we present two simple examples exhibiting some interesting features of the packing spectrum in the two-dimensional case. In §7, we conclude with some extensions of Theorem 5 which follow from Theorem 6, along with some open questions.

2. *Symbolic dynamics*

We begin by restating our theorem in terms of the associated symbolic space. Let Σ denote the symbolic space $\mathcal{D}^{\mathbb{N}}$ under the usual product topology. We let $\Pi : \Sigma \to \Lambda$ denote the natural projection given by

$$\Pi : (\omega_\nu)_{\nu\in\mathbb{N}} \mapsto \left(\sum_{\nu\in\mathbb{N}} \frac{i_j(\nu)}{a_j^\nu}\right)_{j=1}^{d} \quad \text{where } \omega_\nu = (i_j(\nu))_{j=1}^{d} \text{ for } \nu \in \mathbb{N}. \tag{2.1}$$

For each $k = 1, \ldots, d$, we let η_k denote the projection of $\mathcal{D}$ by $\eta_k : (i_j)_{j=1}^{d} \mapsto (i_j)_{j=1}^{d-(k-1)}$ and $\Sigma_k := \eta_k(\mathcal{D})^{\mathbb{N}}$. We then define a projection $\chi_k : \Sigma \mapsto \Sigma_k$, corresponding to $\pi_k : \mathbb{T}^d \to \mathbb{T}^{d-(k-1)}$, by $\chi_k : (\omega_\nu)_{\nu=1}^{\infty} \mapsto (\eta_k(\omega_\nu))_{\nu=1}^{\infty}$. We let $\Pi_k : \Sigma_k \to \pi_k(\Lambda)$ denote the natural projection given by

$$\Pi_k : (\tau_\nu)_{\nu\in\mathbb{N}} \mapsto \left(\sum_{\nu\in\mathbb{N}} \frac{i_j(\nu)}{a_j^\nu}\right)_{j=1}^{d-(k-1)} \quad \text{where } \tau_\nu = (i_j(\nu))_{j=1}^{d-(k-1)} \text{ for } \nu \in \mathbb{N}. \tag{2.2}$$

Note that $\Pi_k \circ \chi_k = \pi_k \circ \Pi$. We let σ denote the left shift on Σ and for each k, σ_k denotes the left shift on Σ_k. Note that $f \circ \Pi = \Pi \circ \sigma$ and for each k, $f_k \circ \Pi_k = \Pi_k \circ \sigma_k$. Given a finite sequence $(\omega_\nu)_{\nu=1}^{n} \in \eta_k(\mathcal{D})$, we let $[\omega_1 \cdots \omega_n]$ denote the cylinder set

$$[\omega_1 \cdots \omega_n] := \{\omega' \in \Sigma_k : \omega'_\nu = \omega_\nu \text{ for } \nu = 1, \ldots, n\}. \tag{2.3}$$

Given $\varphi : \Sigma \to \mathbb{R}$ and $n \in \mathbb{N}$, we define $\mathrm{var}_n(\varphi)$ by

$$\mathrm{var}_n(\varphi) := \sup\{|\varphi(\omega) - \varphi(\tau)| : \omega_\nu = \tau_\nu \text{ for } \nu = 1, \ldots, n\}. \tag{2.4}$$

We also define $A_n(\varphi) : \Sigma \to \mathbb{R}$ to be the map $\omega \mapsto (1/n)\sum_{l=0}^{n-1} \varphi(\sigma^l\omega)$.

We are interested in the space of all Borel probability measures $\mathcal{M}(\Sigma)$ under the weak $*$ topology. Since Σ is compact, and hence the space $C(\Sigma)$ of continuous real-valued functions on Σ is separable, we may choose a countable family of potentials $(\varphi_l)_{l\in\mathbb{N}}$ with norm one, $\|\varphi_l\|_\infty = 1$, for all $l \in \mathbb{N}$, for which sets of the form

$$\left\{\nu \in \mathcal{M}(\Sigma) : \left|\int \varphi_l \, d\nu - \int \varphi_l \, d\mu\right| < \epsilon \text{ for all } l \le L\right\}, \tag{2.5}$$

with $\mu \in \mathcal{M}(\Sigma)$ and $L \in \mathbb{N}$, form a neighbourhood basis of $\mathcal{M}(\Sigma)$.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

For each $n \in \mathbb{N}$, we let $\mathcal{M}_{\sigma^n}(\Sigma)$ denote the set of σ^n-invariant Borel probability measures, let $\mathcal{E}_{\sigma^n}(\Sigma)$ denote the set of $\mu \in \mathcal{M}_{\sigma^n}(\Sigma)$ which are ergodic, with respect to σ^n, and let $\mathcal{B}_{\sigma^n}(\Sigma)$ denote the set of $\mu \in \mathcal{E}_{\sigma^n}(\Sigma)$ which are also Bernoulli.

Given a Borel probability measure $\mu \in \mathcal{M}(\Sigma)$, we define

$$A_n(\mu) := \frac{1}{n} \sum_{l=0}^{n-1} \mu \circ \sigma^{-l}. \tag{2.6}$$

Given $\omega \in \Sigma$, we let $\mathcal{V}(\omega)$ denote the set of all weak $*$ accumulation points of the sequence of measures $(A_n(\delta_\omega))_{n\in\mathbb{N}}$, where δ_ω denotes the Dirac measure concentrated at ω. Given $A \subseteq \mathcal{M}_\sigma(\Sigma)$, we define

$$\begin{aligned} \Gamma(A) &:= \{\omega \in \Omega : \mathcal{V}(\omega) = A\}, \\ \Omega(A) &:= \{\omega \in \Omega : \mathcal{V}(\omega) \subseteq A\}. \end{aligned} \tag{2.7}$$

For each $k \le d$, we define $H^k(\sigma, A)$ by

$$H^k(\sigma, A) := \sup\{h_{\mu\circ\chi_k^{-1}}(\sigma_k) : \mu \in A\}.$$

We shall prove the following theorem, which implies Theorems 5 and 6.

THEOREM 7. *Suppose that A is a non-empty closed convex subset of $\mathcal{M}_\sigma(\Sigma)$. Then,*

$$\dim_{\mathcal{P}}\Pi(\Gamma(A)) = \dim_{\mathcal{P}}\Pi(\Omega(A)) = \frac{H^1(\sigma, A)}{\log a_1} + \sum_{k=2}^{d}\left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}}\right)H^k(\sigma, A).$$

3. *Proof of the lower estimate*

Fix a non-empty closed convex subset $A \subseteq \mathcal{M}(\Sigma, \sigma)$. Take $\zeta > 0$ and choose some $\mu_j \in \mathcal{M}_\sigma(\Sigma)$ for $j = 1, \ldots, d$ such that

$$h_{\mu_j\circ\chi_j^{-1}}(\sigma_j) > H^j(\sigma, A) - \zeta. \tag{3.1}$$

Through a series of lemmas, we shall prove that

$$\dim_{\mathcal{P}}\Pi(\Gamma(A)) \ge \frac{H^1(\sigma, A)}{\log a_1} + \sum_{j=2}^{d}\left(\frac{1}{\log a_j} - \frac{1}{\log a_{j-1}}\right)H^j(\sigma, A). \tag{3.2}$$

To this end, we construct a measure allowing us to apply the following result from geometric measure theory.

PROPOSITION 3.1. *Let $E \subseteq \mathbb{R}^n$ be a Borel set and μ a finite Borel measure. If $\limsup_{r\to 0}(\log \mu(B(x; r))/\log r) \ge s$ for all $x \in E$ and $\mu(E) > 0$, then $\dim_{\mathcal{P}}(E) \ge s$.*

Proof. This follows from [**Fa2**, Proposition 4.9]. □

Let $\lambda_0 := 0$ and for $j = 1, \ldots, d$ we let $\lambda_j := \log a_d/\log a_j$. In order to obtain an optimal lower bound, we shall construct a measure $\mathcal{W}$ which, for infinitly many values of n, behaves like μ_j for the digits from $\lceil\lambda_{j-1}n\rceil + 1$ up to $\lceil\lambda_j n\rceil$, for each $j = 1, \ldots, d$, and use this property to show that $\mathcal{W} \circ \Pi^{-1}$ has the required packing dimension.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

We must also choose $\mathcal{W}$ so that $\mathcal{V}(\omega) = A$ on a set of large $\mathcal{W}$ measure. To do this, we take a sequence of measures $(m_q)_{q\in 2\mathbb{N}}$ in A for which the set of weak $*$ limit points of $(m_q)_{q\in 2\mathbb{N}}$ is precisely the set A. We shall also construct $\mathcal{W}$ so that, along a subsequence of times, $\mathcal{W}$ behaves like $(m_q)_{q\in 2\mathbb{N}}$.

To obtain such a measure, $\mathcal{W}$, we effectively piece together the various invariant measures that $\mathcal{W}$ is required to imitate. In order to carry out this procedure, we must first approximate each of our invariant measures by members of $\bigcup_{n\in\mathbb{N}} \mathcal{B}_{\sigma^n}(\Sigma)$. This allows us to deal with three issues. First, the invariant measures which $\mathcal{W}$ is required to mimic need not be ergodic. Nonetheless, their approximations will be ergodic for some n-shift σ^n, and this allows us to apply both Birkhoff's ergodic theorem and the Shannon–McMillan–Breiman theorem. Secondly, we do not assume King's disjointness condition (see [**Ki**]) and allow our approximate squares to touch at their boundaries. As such, we must ensure that our measure is not too concentrated so that it behaves well under projection by Π. For members of $\bigcup_{n\in\mathbb{N}} \mathcal{B}_{\sigma^n}(\Sigma)$, we may do this simply by tweaking our measure so that it gives each finite word some positive probability. Thirdly, the process of pieceing together measures is greatly simplified by only working with members of $\bigcup_{n\in\mathbb{N}} \mathcal{B}_{\sigma^n}(\Sigma)$. This approximation introduces an error, both in the expected local entropy and expected Birkhoff averages. However, these error terms go to zero as the approximation improves, so, by concatenating increasingly good approximations, we shall obtain a measure which not only behaves well at every given stage, but also gives positive measure to the level set $\Gamma(A)$ and gives an optimal lower bound for the packing dimension of $\Pi(\Gamma(A))$.

Similar techniques appear in the work of Gelfert and Rams [**GR**], Barral and Feng [**BF**], Baek *et al* [**BOS**] and Barreira and Schmeling [**BSch**].

LEMMA 3.1. *For each $j = 1, \ldots, d$ and $q \in 2\mathbb{N}+1$, we may find $k(q) \in \mathbb{N}$ and $\nu_j^q \in \mathcal{B}_{\sigma^{k(q)}}(\Sigma)$ such that for $l = 1, \ldots, q$:*

(i) $|h(\mu_j^q \circ \chi_j^{-1}, \sigma_j) - (1/k(q))h(\nu_j^q \circ \chi_j^{-1}, \sigma_j^{k(q)})| < 1/q$;

(ii) $|\int A_{k(q)}(\varphi_l)\, d\nu_j^q - \int \varphi_l\, d\nu_j| < 1/q$; *and*

(iii) $\nu_j^q([\omega_1 \cdots \omega_{k(q)}]) > 0$ *for all* $(\omega_1, \ldots, \omega_{k(q)}) \in \mathcal{D}^{k(q)}$.

Proof. Given $j \in \mathbb{N}$ and $k \in \mathbb{N}$, we let μ_j^k denote the unique member of $\mathcal{B}_{\sigma^k}(\Sigma)$ which agrees with μ_j on cylinders of length k. So, for all $(\tau_1, \ldots, \tau_k) \in \eta_j(\mathcal{D})^k$, we have

$$\mu_j^k \circ \chi_j^{-1}([\tau_1 \cdots \tau_k]) = \mu_a \circ \chi_j^{-1}([\tau_1 \cdots \tau_k]).$$

Now, by the Kolmogorov–Sinai theorem [**Wa**, Theorem 4.18], we have

$$h(\mu_j \circ \chi_j^{-1}, \sigma_j) = -\lim_{k\to\infty} \frac{1}{k} \sum_{(\tau_1,\ldots,\tau_k)\in\eta(\mathcal{D})^k} \mu_j \circ \chi_j^{-1}([\omega_1, \ldots, \omega_k]) \log \mu_j \circ \chi_j^{-1}([\omega_1, \ldots, \omega_k]).$$

Equivalently,

$$h(\mu_j \circ \chi_j^{-1}, \sigma_j) = \lim_{k\to\infty} \frac{1}{k} h(\mu_j^k \circ \chi_j^{-1}, \sigma^k).$$

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Since each μ_j is σ invariant, we have $\int A_k(\varphi_l)\,d\mu_j = \int \varphi_l\,d\mu_j$ for $l = 1, \ldots, k$ and, as μ_j and μ_j^k agree on cylinders of length k, we have

$$\left|\int A_k(\varphi_l)\,d\mu_j^k - \int A_k(\varphi_l)\,d\mu_j\right| \le \frac{1}{k}\sum_{n=0}^{k-1} \mathrm{var}_n(\varphi_l).$$

Moreover, since each φ_l is continuous, $\mathrm{var}_k(\varphi_l) \to 0$ (and hence $(1/k)\sum_{n=0}^{k-1}\mathrm{var}_n(\varphi_l) \to 0$) as $k \to \infty$. Thus, for each $l = 1, \ldots, q$,

$$\lim_{k\to\infty} \int A_k(\varphi_l)\,d\mu_j^k = \int \varphi_l\,d\mu_j.$$

Thus, taking $\nu_j^q = \mu_j^{k(q)}$, for each j, a sufficiently large $k(q) \in \mathbb{N}$ gives (i) and (ii). By slightly adjusting ν_j^q, we may ensure that (i), (ii) and (iii) hold. □

LEMMA 3.2. *For all $q \in 2\mathbb{N}$, we may find $k(q) \in \mathbb{N}$ and $\tilde{m}_q \in \mathcal{B}_{\sigma^{k(q)}}(\Sigma)$ such that for $l = 1, \ldots, q$:*
(i) $|\int A_{k(q)}(\varphi_l)\,d\tilde{m}_q - \int \varphi_l\,dm_q| < 1/q$; *and*
(ii) $\tilde{m}_q([\omega_1 \cdots \omega_{k(q)}]) > 0$ *for all* $(\omega_1, \ldots, \omega_{k(q)}) \in \mathcal{D}^{k(q)}$.

Proof. Essentially the same as for Lemma 3.1. □

Choose $\delta_q > 0$ for each $q \in \mathbb{N}$ so that $\prod_{q\in\mathbb{N}}(1-\delta_q) > 0$.

LEMMA 3.3. *For each $j = 1, \ldots, d$ and $q \in 2\mathbb{N}+1$, we may find $N(q) \in \mathbb{N}$ and a subset $S_j^q \subseteq \Sigma$ with $\nu_j^q(S_j^q) > 1-\delta_q$ and such that for all $\omega \in S_j^q$ and $n \ge N(q)$, and all $l = 1, \ldots, q$, we have:*
(i) $|(1/nk(q))\sum_{r=0}^{nk(q)-1}\varphi_l(\sigma^r\omega) - \int \varphi_l\,d\mu_j| < 1/q$;
(ii) $|(1/nk(q))\log \nu_j^q \circ \chi_j^{-1}([\eta_j(\omega_1)\cdots\eta_j(\omega_{nk(q)})]) + h(\mu_j \circ \chi_j^{-1}, \sigma_j)| < 1/q$; *and*
(iii) $\{d \in \mathcal{D} : \omega_r = d \text{ for some } r \le N(q)\} = \mathcal{D}$.

Proof. Given $q \in \mathbb{N}$, we may apply the Birkhoff ergodic theorem and the Shannon–Breiman–McMillan theorem to $\nu_j^q \circ \chi_j^{-1} \in \mathcal{B}_{\sigma_j^{k(q)}}(\Sigma_j) \subseteq \mathcal{E}_{\sigma_j^{k(q)}}(\Sigma_j)$ to obtain

$$\lim_{n\to\infty}\frac{1}{nk(q)}\sum_{r=0}^{nk(q)-1}\varphi_l(\sigma^r\omega) = \lim_{n\to\infty}\frac{1}{n}\sum_{r=0}^{n-1}A_{k(q)}(\varphi_l)(\sigma^{rk(q)}\omega) = \int \varphi_l\,d\nu_j^q, \tag{3.3}$$

$$\lim_{n\to\infty}\frac{1}{n}\log \mu_j^q \circ \chi_j^{-1}([\eta_j(\omega_1)\cdots\eta(\omega_{nk(q)})]) = -h(\mu_j^q \circ \chi_j^{-1}, \sigma_j^{k(q)}) \tag{3.4}$$

for ν_j^q almost every $\omega \in \Sigma_j$.

By Egorov's theorem, we may choose subsets $S_j^q \subseteq \Sigma$ with $\nu_j^q(S_j^q) > 1-\delta_q$ so that the convergences in (3.3) and (3.4) are uniform on S_j^q. Thus, by Lemma 3.1, we choose $N(q) \in \mathbb{N}$ so that, for all $n \ge N(q)$ and all $\omega \in S_j^q$, we have

$$\left|\frac{1}{nk(q)}\sum_{r=0}^{nk(q)-1}\varphi_l(\sigma^r\omega) - \alpha\right| < \frac{1}{q}, \tag{3.5}$$

$$\left|\frac{1}{nk(q)}\log \nu_a^q \circ \chi_j^{-1}([\eta(\omega_1)\cdots\eta(\omega_{nk(q)})]) + h(\mu_j \circ \chi_j^{-1}, \sigma_j)\right| < \frac{1}{q}. \tag{3.6}$$

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

In light of condition (iii), for ν_j^q almost every $\omega \in \Sigma$, we have $\{d \in \mathcal{D} : \omega_r = d \text{ for some } r \in \mathbb{N}\} = \mathcal{D}$. Equivalently, for ν_j^q almost every $\omega \in \Sigma$, there exists some $M(\omega) \in \mathbb{N}$ for which $\{d \in \mathcal{D} : \omega_r = d \text{ for some } r \le M(\omega)\} = \mathcal{D}$. Thus, by moving to a subset of S_j^q of large ν_j^q measure, and increasing $N(q)$, if necessary, we may assume that for all $\omega \in S_j^q$ and all $n \ge N(q)$, we have

$$\{d \in \mathcal{D} : \omega_l = d \text{ for some } l \le N(q)\} = \mathcal{D}. \tag{3.7}$$

□

LEMMA 3.4. *For each $q \in 2\mathbb{N}$, we may find $N(q) \in \mathbb{N}$ and a subset $S^q \subseteq \Sigma$ with $\tilde{m}^q(S^q) > 1 - \delta_q$ and such that for all $\omega \in S^q$ and $n \ge N(q)$, and all $l = 1, \ldots, q$, we have:*

(i) $|(1/nk(q))\sum_{r=0}^{nk(q)-1} \varphi_l(\sigma^r\omega) - \int \varphi_l \, dm_q| < 1/q$; *and*

(ii) $\{d \in \mathcal{D} : \omega_r = d \text{ for some } r \le N(q)\} = \mathcal{D}$.

Proof. Essentially the same as for Lemma 3.3. □

We shall now construct a probability measure $\mathcal{W}$ on Σ. To do this, we first define a rapidly increasing sequence of natural numbers $(\gamma_q)_{q\in\mathbb{N}}$ as follows. Let $\gamma_0 := 0$ and for each $q \ge 1$ take some $\gamma_q > (q+1)\prod_{j=1}^d (\lambda_j - \lambda_{j-1})^{-1}(\prod_{r=1}^{q+2} N(r)k(r) + \gamma_{q-1})$ so that $\gamma_q - \gamma_{q-1}$ is divisible by $k(q)$. For each $j = 1, \ldots, d$, we define sequences of natural numbers $(\vartheta_q^j)_{q\in 2\mathbb{N}+1}$ by letting ϑ_q^j denote the greatest integer which is divisible by $k(q)$ and does not exceed $\lambda_k \gamma_q$. For simplicity, we also let $\vartheta_q^0 := \gamma_{q-1}$.

We define a measure $\mathcal{W}$ on Σ by first defining $\mathcal{W}$ on cylinders of length γ_{2Q} for some $Q \in \mathbb{N}$, and then extending $\mathcal{W}$ to a Borel probability measure via the Daniell–Kolmogorov consistency theorem (see [**Wa**, §0.5]). Given a cylinder $[\omega_1 \cdots \omega_{\gamma_{2Q}}]$ of length γ_{2Q}, we let

$$\begin{aligned}&\mathcal{W}([\omega_1 \cdots \omega_{\gamma_{2Q}}]) \\ &:= \prod_{q=1}^{Q}\left(\prod_{j=1}^{d} \nu_j^{2q-1}([\omega_{\vartheta_{2q-1}^{j-1}+1} \cdots \omega_{\vartheta_{2q-1}^{j}}]) \times \tilde{m}^{2q}([\omega_{\gamma_{2q-1}+1} \cdots, \omega_{\gamma_{2q}}])\right).\end{aligned} \tag{3.8}$$

Define $S \subseteq \Sigma$ by

$$\begin{aligned}S := \bigcap_{q=1}^{\infty}\Bigg(\bigcap_{j=1}^{d}\{\omega \in \Sigma : [\omega_{\vartheta_{2q-1}^{j-1}+1} \cdots \omega_{\vartheta_{2q-1}^{j}}] \cap S_j^{2q-1} \neq \emptyset\} \\ \cap \{\omega \in \Sigma : [\omega_{\gamma_{2q-1}+1} \cdots, \omega_{\gamma_{2q}}] \cap S^{2q} \neq \emptyset\}\Bigg).\end{aligned} \tag{3.9}$$

LEMMA 3.5. $\mathcal{W}(S) > 0$.

Proof. The proof is as follows.

$$\mathcal{W}(S) \ge \prod_{q=1}^{\infty}\left(\left(\prod_{j=1}^{d} \nu_j^{2q-1}(S_j^{2q-1})\right)\tilde{m}^{2q}(S^{2q})\right) > \prod_{q=1}^{\infty}(1-\delta_q)^d > 0. \tag{3.10}$$

□

LEMMA 3.6. *For all $\omega \in S$, $\mathcal{V}(\omega) \subseteq A$.*

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Proof. Choose $\omega \in S$ and fix $Q \in \mathbb{N}$ and $\epsilon > 0$. For each $q \in 2\mathbb{N}$ with $q \geq Q$, take $\tau^q \in [\omega_{\gamma_{q-1}+1} \cdots \omega_{\gamma_q}] \cap S^q$, which is non-empty since $\omega \in S$. By Lemma 3.4, for all $n \geq k(q)N(q)$,

$$\left| \sum_{r=0}^{n-1} \varphi_l(\sigma^r \tau^q) - n \int \varphi_l \, dm_q \right| < \frac{n}{q} + k(q). \tag{3.11}$$

Hence, for all $N(q)k(q) \leq n \leq \gamma_q - \gamma_{q-1}$,

$$\left| \sum_{r=\gamma_{q-1}}^{\gamma_{q-1}+n-1} \varphi_l(\sigma^r \omega) - n \int \varphi_l \, dm_q \right| < \sum_{r=0}^{n} \mathrm{var}_r(\varphi_l) + \frac{n}{q} + k(q). \tag{3.12}$$

In a similar way, we can show that for all $q \in 2\mathbb{N} - 1$, with $q \geq Q$, $j = 1, \ldots, d$, and all $N(q)k(q) \leq n \leq \vartheta_q^{j-1} - \vartheta_{q-1}^{j}$, we have

$$\left| \sum_{r=\vartheta_q^{j-1}}^{\vartheta_q^{j-1}+n-1} \varphi_l(\sigma^r \omega) - n \int \varphi_l \, d\mu_j \right| < \sum_{r=0}^{n} \mathrm{var}_r(\varphi_l) + \frac{n}{q} + k(q). \tag{3.13}$$

Moreover, given any $n, k, q \in \mathbb{N}$, we automatically have

$$\left| \sum_{r=k}^{k+n} \varphi_l(\sigma^r \omega) - n \int \varphi_l \, dm_q \right| < n. \tag{3.14}$$

Suppose $\gamma_{2q} < N < \gamma_{2q+2}$, where $2q - 2 \geq Q$. Now consider the sum $\sum_{r=0}^{N-1} \varphi_l(\sigma^r \omega)$, for $l \leq Q$. First, break the sum down as follows:

$$\sum_{r=0}^{N-1} \varphi_l(\sigma^r \omega) = \sum_{r=0}^{\gamma_{2q}-1} \varphi_l(\sigma^r \omega) + \sum_{r=\gamma_{2q}}^{N-1} \varphi_l(\sigma^r \omega). \tag{3.15}$$

To deal with the first summand, $\sum_{r=0}^{\gamma_{2q}-1} \varphi_l(\sigma^r \omega)$, we write

$$\sum_{r=0}^{\gamma_{2q}-1} \varphi_l(\sigma^r \omega) = \underbrace{\sum_{r=0}^{\gamma_{2q-2}-1} \varphi_l(\sigma^r \omega)}_{*} + \underbrace{\sum_{j=1}^{d} \sum_{r=\vartheta_{2q-1}^{j-1}}^{\vartheta_{2q-1}^{j}-1} \varphi_l(\sigma^r \omega)}_{**} + \underbrace{\sum_{r=\gamma_{2q-1}}^{\gamma_{2q}-1} \varphi_l(\sigma^r \omega)}_{***}. \tag{3.16}$$

To part $*$, we apply (3.14), whilst to each of the parts labelled $**$, we apply (3.13), and to the part labelled $***$, we apply (3.12).

For the second summand, $\sum_{r=\gamma_{2q}}^{N-1} \varphi_l(\sigma^r \omega)$, there are two cases. Either we have $N \leq \gamma_{2q+1}$ or $N > \gamma_{2q+1}$. In the former case, we have

$$\sum_{r=\gamma_{2q}}^{N-1} \varphi_l(\sigma^r \omega) = \sum_{j=1}^{J} \underbrace{\sum_{r=\vartheta_{2q}^{j-1}}^{\vartheta_{2q}^{j}-1} \varphi_l(\sigma^r \omega)}_{**} + \underbrace{\sum_{r=\vartheta_{2q}^{J}}^{N-1} \varphi_l(\sigma^r \omega)}_{\dagger}, \tag{3.17}$$

where J is the greatest $j \in \{1, \ldots, d\}$ such that $J < N - 1$. To parts labelled $**$, we again apply (3.13), and to the part labelled (†) we either apply (3.14) or (3.13),

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

depending on whether $N - \vartheta_{2q}^J - 1 < \max_{j\in\{2q,2q+1\}}\{k(j)N(j)\}$ or $N - \vartheta_{2q}^J - 1 \geq \max_{j\in\{2q,2q+1\}}\{k(j)N(j)\}$. In the latter case, we have

$$\sum_{r=\gamma_{2q}}^{N-1}\varphi_l(\sigma^r\omega) = \sum_{j=1}^{d}\underbrace{\sum_{r=\vartheta_{2q}^{j-1}}^{\vartheta_{2q}^{j}-1}\varphi_l(\sigma^r\omega)}_{**} + \underbrace{\sum_{r=\gamma_{2q+1}}^{N-1}\varphi_l(\sigma^r\omega)}_{\dagger\dagger}. \tag{3.18}$$

Again, to the parts labelled $**$, we apply (3.13), and to the part labelled (††), we either apply (3.14) or (3.12), depending on whether $N - \gamma_{2q+1} - 1 < \max_{j\in\{2q,2q+1\}}\{k(j)N(j)\}$ or $N - \gamma_{2q+1} - 1 \geq \max_{j\in\{2q,2q+1\}}\{k(j)N(j)\}$.

Thus, by combining (3.16), (3.17) and (3.18), in each case we see that there exist $\beta_{2q}^N, \beta_{2q+2}^N, \lambda_1^N, \ldots, \lambda_d^N \in [0,1]$ which sum to one, depending solely on N and not on $l = 1, \ldots, Q$, for which we have

$$\begin{aligned}
&\left|\sum_{r=0}^{N-1}\varphi_l(\sigma^r\omega) - \sum_{j\in\{2q,2q+2\}}N\beta_j^N\int\varphi_l\,dm_j - \sum_{j=1}^{d}N\lambda_j^N\int\varphi_l\,d\mu_j\right| \\
&\quad< (2d+1)\left(\sum_{r=0}^{N}\mathrm{var}_r(\varphi_l) + \frac{N}{q-1} + \max_{j\in\{0,1,2,3\}}\{k(2q+j)\}\right) \\
&\qquad+ \left(\gamma_{2q-2} + \max_{j\in\{2q,2q+1\}}\{k(j)N(j)\}\right) \\
&\quad< o(N) + o(\gamma_{2q}) \leq o(N),
\end{aligned}$$

where we use the continuity of each φ_l together with the definition of $(\gamma_q)_{q\in\mathbb{N}}$ to obtain the last line. Moreover, since A is convex, for each such N, the measure $\rho_N := \sum_{j\in\{2q,2q+2\}}N\beta_j^N m_j + \sum_{j=1}^{d}\lambda_j^N\mu_j$ is a member of A. Hence, for all sufficiently large N, we have

$$\left|\int\varphi_l\,dA_N(\delta_\omega) - \int\varphi_l\,d\rho_N\right| < \epsilon, \tag{3.19}$$

for $l = 1, \ldots, Q$. Since A is also closed, it follows that every weak $*$ accumulation point of the sequence $(A_N(\delta_\omega))_{N\in\mathbb{N}}$ is a member of A. □

LEMMA 3.7. *For all $\omega \in S$, $A \subseteq \mathcal{V}(\omega)$.*

Proof. Take $\omega \in S, \alpha \in A$. Since the set of accumulation points of $(m_q)_{q\in2\mathbb{N}}$ is equal to A, we may extract a subsequence $(m_{q_j})_{j\in\mathbb{N}}$ converging to ρ. Now choose $\epsilon > 0$ and choose Q so large that, for j with $q_j \geq Q$,

$$\left|\int\varphi_l\,dm_{q_j} - \int\varphi_l\,d\rho\right| < \epsilon. \tag{3.20}$$

As in (3.12), we see that for all j with $q_j \geq Q, l = 1, \ldots, Q$,

$$\left|\sum_{r=\gamma_{q_j-1}}^{\gamma_{q_j}-1}\varphi_l(\sigma^r\omega) - (\gamma_{q_j} - \gamma_{q_j-1})\int\varphi_l\,dm_{q_j}\right| \tag{3.21}$$

$$< \sum_{r=0}^{\gamma_{q_j}-\gamma_{q_j-1}}\mathrm{var}_r(\varphi_l) + \frac{\gamma_{q_j} - \gamma_{q_j-1}}{q_j} + k(q_j). \tag{3.22}$$

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Hence, for all $q_j \geq Q, l = 1, \ldots, Q$,

$$\left| \sum_{r=0}^{\gamma_{q_j}-1} \varphi_l(\sigma^r \omega) - \gamma_{q_j} \int \varphi_l \, dm_{q_j} \right| \tag{3.23}$$

$$< \gamma_{q_j-1} + \sum_{r=0}^{\gamma_{q_j}-\gamma_{q_j-1}} \mathrm{var}_r(\varphi_l) + \frac{\gamma_{q_j} - \gamma_{q_j-1}}{q_j} + k(q_j). \tag{3.24}$$

Thus, by the definition of $(\gamma_q)_{q\in\mathbb{N}}$, and the fact that each φ_l is continuous, we have

$$\left| \int \varphi_l \, dA_{\gamma_{q_j}}(\delta_\omega) - \int \varphi_l \, d\rho \right| < 2\epsilon. \tag{3.25}$$

It follows that $A_{\gamma_{q_j}}(\delta_\omega) \to \rho$ as $j \to \infty$. □

LEMMA 3.8. $S \subseteq \Gamma(A)$.

Proof. Combine Lemmas 3.6 and 3.7. □

For each $q \in 2\mathbb{N}+1$, we define the qth approximate square $B_q^L(\omega)$ to be the set

$$B_q^L(\omega) := \{\omega' \in \Sigma : \text{for } j = 1, \ldots, d,\ \eta_j(\omega'_\nu) = \eta_j(\omega_\nu) \text{ for } \nu = 1, \ldots, \vartheta_q^j\}.$$

LEMMA 3.9. *For all $\omega \in S$,*

$$\limsup_{q\to\infty} \frac{1}{\gamma_q} \log \mathcal{W}(B_q^L(\omega)) \leq -\lambda_1 h(\mu_1, \sigma) - \sum_{j=1}^{d} (\lambda_j - \lambda_{j-1}) h(\mu_j \circ \chi_j^{-1}, \sigma_j).$$

Proof. Take $\omega \in S$ and $q \in 2\mathbb{N}+1$ and $j \in \{1, \ldots, d\}$. Since $\vartheta_q^j - \vartheta_q^{j-1} \geq N(q)$ and $[\omega_{\gamma_{q-1}+1} \cdots \omega_{\vartheta_q}] \cap S_j^q \neq \emptyset$, it follows from the definition of $\mathcal{W}$ and $(\gamma_q)_{q\in\mathbb{N}}$ together with Lemma 3.3 that

$$\begin{aligned}
&\log \mathcal{W}(\sigma^{-\vartheta_q^{j-1}}[\eta_j(\omega_{\vartheta_q^{j-1}+1}) \cdots \eta_j(\omega_{\vartheta_q^j})]) \\
&= \log \nu_j^q([\eta_j(\omega_{\vartheta_q^{j-1}+1}) \cdots \eta_j(\omega_{\vartheta_q^j})]) \\
&\leq -(\vartheta_q^j - \vartheta_q^{j-1}) h(\mu_j^q \circ \chi_j^{-1}, \sigma_j) + \frac{\vartheta_q^j - \vartheta_q^{j-1}}{q} \\
&\leq -((\lambda_j - \lambda_{j-1})\gamma_q - k(q) - \gamma_{q-1}) h(\mu_j^q \circ \chi_j^{-1}, \sigma_j) + \frac{\gamma_q}{q} \\
&\leq -(\lambda_j - \lambda_{j-1})\gamma_q h(\mu_j^q \circ \chi_j^{-1}, \sigma_j) + o(\gamma_q).
\end{aligned} \tag{3.26}$$

By the definition of $\mathcal{W}$, each of the cylinders $\sigma^{-\vartheta_q^{j-1}}[\eta_j(\omega_{\vartheta_q^{j-1}+1}) \cdots \eta_j(\omega_{\vartheta_q^j})]$, with $j = 1, \ldots, d$, is independent with respect to $\mathcal{W}$. Thus, letting $q \to \infty$, we have

$$\limsup_{q\to\infty} \frac{1}{\gamma_q} \log \mathcal{W}\left(\bigcap_{j=1}^{d} \sigma^{-\vartheta_q^{j-1}}[\eta_j(\omega_{\vartheta_q^{j-1}+1}) \cdots \eta_j(\omega_{\vartheta_q^j})] \right)$$

$$\leq -\lambda_1 h(\mu_1, \sigma) - \sum_{j=1}^{d} (\lambda_j - \lambda_{j-1}) h(\mu_j \circ \chi_j^{-1}, \sigma_j). \tag{3.27}$$

Since $B_q^L(\omega) \subseteq \bigcap_{j=1}^d \sigma^{-\vartheta_q^{j-1}}[\eta_j(\omega_{\vartheta_q^{j-1}+1}) \cdots \eta_j(\omega_{\vartheta_q^j})]$, the lemma follows. □

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

The following lemma allows us to deal with the fact that our approximate squares may meet at their boundaries.

LEMMA 3.10. *For all $\omega \in S$ and $q \in \mathbb{N}$, we have*

$$B\left(\Pi(\omega); \min_{j\in\{1,\dots,d\}}\{a_j^{-\vartheta_q^j-\max\{N(q),N(q+1)\}}\}\right)\cap\Lambda \subseteq \Pi(B_q^L(\omega)).$$

Proof. Fix $\omega = ((i_\nu^j)_{j=1}^d)_{\nu\in\mathbb{N}} \in S$ and let $M(q) := \max\{N(q), N(q+1)\}$. Clearly, it suffices to show that for each $j = 1, \dots, d$,

$$B\left(\Pi_j(\omega); \min_{j\in\{1,\dots,d\}}\{a_j^{-\vartheta_q^j-M(q)}\}\right)\cap\Lambda \subseteq \Pi_j(B_q^L(\omega)). \tag{3.28}$$

Take $j \in \{1, \dots, d\}$ and let $x = \Pi_j(\omega)$. We may divide $\Pi_j(B_q^L(\omega))$ into $\#\eta_j(\mathcal{D})^{M(q)}$ intervals of width $a_j^{-\vartheta_q^j-M(q)}$ with disjoint interiors, each corresponding to a possible string of digits $i'_{\vartheta_q^j+1}\cdots i'_{\vartheta_q^j+M(q)}$ for $\tau = ((\tilde{i}_\nu^j)_{j=1}^d)_{\nu\in\mathbb{N}} \in B_q^L(\omega)$. Since $\omega \in S$, we have

$$[\omega_{\vartheta_q^j+1}\cdots\omega_{\vartheta_q^j+N(q)}]\cap S_{j+1}^q \neq \emptyset \quad \text{for } j = 1, \dots, d-1$$

and

$$[\omega_{\vartheta_q^d+1}\cdots\omega_{\vartheta_q^d+N(q+1)}]\cap S_d^{q+1} \neq \emptyset.$$

Thus, by Lemma 3.3 in the first case and Lemma 3.4,

$$\mathcal{D} = \{d \in \mathcal{D} : \omega_l = d \text{ for some } \vartheta_q^j < l \le \vartheta_q^j + M(q)\}.$$

It follows from $\eta_j(\mathcal{D}) = \{d \in \eta_j(\mathcal{D}) : i_\nu^j = d \text{ for some } \vartheta_q^j < l \le \vartheta_q^j + N(q)\}$ that x is in neither the far left nor the far right interval of $\Pi_j(B_q^L(\omega))$, for, in either case, $(i_\nu^j)_{\nu=\vartheta_q^j+1}^{\vartheta_q^j+N(q)}$ would be a constant sequence. Since $\#\eta_j(\mathcal{D}) > 1$ it follows that x is a distance at least $a_j^{-\vartheta_q^j-M(q)}$ away from any point y such that y is an a_j-ary digit expansion $(\tilde{i}_\nu^j)_{\nu\in\mathbb{N}}$ with $\tilde{i}_\nu^j \neq i_\nu^j$ for some $\nu \le \vartheta_q^j$. Thus, (3.28) holds. □

Let $\mathcal{M} := \mathcal{W}\circ\chi^{-1}$ denote the pushdown of $\mathcal{W}$ onto Λ.

LEMMA 3.11. *For all $\omega \in S$,*

$$\limsup_{r\to 0}\frac{\log\mathcal{M}(B(\chi(\omega);r))}{\log r} \ge -\frac{h(\mu_1,\sigma)}{\log a_1} + \sum_{j=2}^{d}\left(\frac{1}{\log a_j}-\frac{1}{\log a_{j-1}}\right)h(\mu_j\circ\chi_j^{-1},\sigma_j).$$

Proof. Recall that for each $j = 1, \dots, d$, we defined $\lambda_j := \log a_d/\log a_j$, and for each $q \in 2\mathbb{N}+1$, we have $\vartheta_q^j < \lambda_j\gamma_q$, and so

$$\min_{j\in\{1,\dots,d\}}\{a_j^{-\vartheta_q^j-\max\{N(q),N(q+1)\}}\} \ge a_d^{-\gamma_q-\max\{N(q),N(q+1)\}/\lambda_1}. \tag{3.29}$$

Choose $\omega \in S$. By Lemma 3.10,

$$B(\Pi(\omega); a_d^{-\gamma_q-\max\{N(q),N(q+1)\}/\lambda_1})\cap\Lambda \subseteq \Pi(B_q^L(\omega)). \tag{3.30}$$

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Hence,

$$\mathcal{M}(B(\Pi(\omega); a_d^{-\gamma_q - \max\{N(q), N(q+1)\}/\lambda_1})) \leq \mathcal{W}(B_q^L(\omega)). \tag{3.31}$$

It follows from Lemma 3.9 that

$$\limsup_{q\to\infty} \frac{1}{\gamma_q} \log \mathcal{M}(B(\Pi(\omega); a_d^{-\gamma_q - \max\{N(q), N(q+1)\}/\lambda_1}))$$
$$\leq -\lambda_1 h(\mu_1, \sigma) - \sum_{j=2}^{d} (\lambda_j - \lambda_{j-1}) h(\mu_j \circ \chi_j^{-1}, \sigma_j). \tag{3.32}$$

Thus, noting that $\max\{N(q), N(q+1)\} = o(q)$, by the definition of $(\gamma_q)_{q\in\mathbb{N}}$, we have

$$\limsup_{q\to\infty} \frac{\log \mathcal{M}(B(\chi(\omega); a_d^{-\gamma_q - \max\{N(q), N(q+1)\}/\lambda_1}))}{\log a_d^{-\gamma_q - \max\{N(q), N(q+1)\}/\lambda_1}}$$
$$\geq -\frac{h(\mu_1, \sigma)}{\log a_1} + \sum_{j=2}^{d} \left(\frac{1}{\log a_j} - \frac{1}{\log a_{j-1}} \right) h(\mu_j \circ \chi_j^{-1}, \sigma_j). \tag{3.33}$$

□

Since $\mathcal{M}(\Pi(S)) \geq \mathcal{W}(S) > 0$, we may combine Proposition 3.1 with Lemma 3.11, to see that

$$\dim_{\mathcal{P}}(\Pi(S)) \geq -\frac{h(\mu_1, \sigma)}{\log a_1} + \sum_{j=2}^{d} \left(\frac{1}{\log a_j} - \frac{1}{\log a_{j-1}} \right) h(\mu_j \circ \chi_j^{-1}, \sigma_j). \tag{3.34}$$

Thus, by Lemma 3.8 and our choice of μ_j (3.1), we have

$$\dim_{\mathcal{P}} \Pi(\Gamma(A)) \geq \frac{H^1(\sigma, A)}{\log a_1} + \sum_{j=2}^{d} \left(\frac{1}{\log a_j} - \frac{1}{\log a_{j-1}} \right) H^j(\sigma, A) - \zeta \frac{d}{\log a_1}. \tag{3.35}$$

By letting $\zeta \to 0$, this completes the proof of the lower bound.

4. *Proof of the upper bound*

Take $A \subseteq \mathcal{M}_\sigma(\Sigma)$. Recall that for each $q \in \mathbb{N}$, we defined

$$U(A, q) := \left\{ \mu \in \mathcal{M}(\Sigma) : \exists \nu \in A, \text{ for all } l \leq q, \left| \int \varphi_l \, d\mu - \int \varphi_l \, d\nu \right| \leq \frac{1}{q} \right\}. \tag{4.1}$$

Given $N \in \mathbb{N}$, we let

$$\Omega(A, N, q) := \{\omega \in \Sigma : \text{for all } n \geq N, \, A_n(\delta_\omega) \in U(A, q)\}. \tag{4.2}$$

Note that for each $q \in \mathbb{N}$, $\{\Pi(\Omega(A, N, q))\}_{N\in\mathbb{N}}$ is a countable cover of $\Pi(\Omega(A))$. As such, we shall give an estimate for the upper box dimension of the sets $\Pi(\Omega(A, N, q))$ before applying the following reformulation of the notion of packing dimension.

PROPOSITION 4.1. *Given $E \subseteq \mathbb{R}^n$, we have*

$$\dim_{\mathcal{P}} E = \inf \left\{ \sup_{n\in\mathbb{N}} \overline{\dim}_B E_n : E \subseteq \bigcup_{n\in\mathbb{N}} E_n \right\},$$

where the infimum is taken over all countable covers $\{E_n\}_{n\in\mathbb{N}}$ of E.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

The above formula is equivalent to the usual definition of packing dimension in terms of s-dimensional packing measures (see [**Mat**, §5.9, Theorem 5.11]).

Recall that for each $j = 1, \ldots, d$, $\lambda_j := \log a_d / \log a_j$. We also let $\lambda_0 := 0$. Given $n \in \mathbb{N}$, we define $\mathcal{A}_n^j$ to be the set of all $(\tau_{\lceil \lambda_{j-1} n \rceil + 1}, \ldots, \tau_{\lceil \lambda_j n \rceil}) \in \eta_j(\mathcal{D})^{\lceil \lambda_j n \rceil - \lceil \lambda_{j-1} n \rceil}$ satisfying

$$\chi_j(\Omega(A, N, q)) \cap \sigma_j^{-\lceil \lambda_{j-1} n \rceil}[\tau_{\lceil \lambda_{j-1} n \rceil + 1}, \ldots, \tau_{\lceil \lambda_j n \rceil}] \neq \emptyset. \tag{4.3}$$

LEMMA 4.1. *Given $N, q \in \mathbb{N}$, we have*

$$\overline{\dim}_B \Pi(\Omega(A, N, q)) \leq \sum_{j=1}^{d} \limsup_{n \to \infty} \frac{\log \#\mathcal{A}_n^j}{n \log a_d}.$$

Proof. Given $r > 0$, we let $N(r)$ denote the minimal number of balls of radius r required to cover $\Pi(\Omega(A, N, q))$ so that

$$\overline{\dim}_B \Pi(\Omega(A, N, q)) = \limsup_{r \to 0} \frac{\log N(r)}{\log(1/r)}.$$

For each n_r, we take $n_r \in \mathbb{N}$ so that $a_d^{-n_r} < r \leq a_d^{-n_r+1}$. Given $\kappa := (\kappa_j)_{j=1}^d$, where $\kappa_j = (\tau^j_{\lceil \lambda_{j-1} n \rceil + 1}, \ldots, \tau^j_{\lceil \lambda_j n \rceil}) \in \mathcal{A}_{n_r}^j$, we let $B(\kappa)$ denote the approximate square

$$B(\kappa) := \Pi\left(\bigcap_{j=1}^{d} \sigma^{-\lceil \lambda_{j-1} n \rceil} \chi_j^{-1}([\tau_{\lceil \lambda_{j-1} n \rceil + 1}, \ldots, \tau_{\lceil \lambda_j n \rceil}]) \right). \tag{4.4}$$

It follows from the definition of Π that each $B(\kappa)$ has diameter no greater than $a_j^{-\lceil \lambda_j n_r \rceil} \leq a_d^{-n_r}$. Moreover,

$$\Pi(\Omega(A, N, q)) \subseteq \bigcup_{\kappa \in \prod_{j=1}^d \mathcal{A}_n^j} B(\kappa). \tag{4.5}$$

Thus,

$$N(r) \leq \#\left\{ B(\kappa) : \kappa \in \prod_{j=1}^{d} \mathcal{A}_{n_r}^j \right\} = \prod_{j=1}^{d} \#\mathcal{A}_{n_r}^j. \tag{4.6}$$

Hence, since $r \leq a_d^{-n_r+1}$,

$$\begin{aligned} \limsup_{r \to 0} \frac{\log N(r)}{\log(1/r)} &\leq \limsup_{r \to 0} \frac{\sum_{j=1}^d \log \#\mathcal{A}_{n_r}^j}{n_r \log a_d} \frac{n_r}{n_r - 1} \\ &\leq \limsup_{n \to \infty} \frac{\sum_{j=1}^d \log \#\mathcal{A}_n^j}{n \log a_d} \\ &\leq \sum_{j=1}^{d} \limsup_{n \to \infty} \frac{\log \#\mathcal{A}_n^j}{n \log a_d}. \end{aligned} \tag{4.7}$$

□

Recall that given $\nu \in \mathcal{M}_{\sigma^k}(\Sigma)$ for each $k \in \mathbb{N}$, we let $A_k(\nu) := (1/k) \sum_{l=0}^{k-1} \nu \circ \sigma^{-l}$.

LEMMA 4.2. *If $\mu = A_k(\nu)$ for some $\nu \in \mathcal{E}_{\sigma^k}(\Sigma)$, then for each $j = 1, \ldots, d$ and all $\varphi \in C(\Sigma)$:*

(i) $\mu \circ \chi_j^{-1} \in \mathcal{E}_{\sigma_j}(\Sigma_j)$;

(ii) $h(\mu \circ \chi_j^{-1}, \sigma_j) = 1/k h(\nu \circ \chi_j^{-1}, \sigma_j^k)$; *and*

(iii) $\int \varphi \, d\mu = \int A_k(\varphi) \, d\nu$.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Proof. By noting that $A_k(\nu) \circ \chi_j^{-1} = \sum_{r=1}^{k-1} \nu \circ \chi_j^{-1} \circ \sigma_j^{-r}$, we see that Lemma 4.2 follows from lemma in [**JJOP**, Lemma 2]. □

For each $l \in \mathbb{N}$, we define

$$H^j(A, l) := \sup\{h(\mu \circ \chi_j^{-1}, \sigma_j) : \mu \in U(A, l)\}. \tag{4.8}$$

Define a constant $L \in (0, 1)$ by

$$L := \min_{j\in\{1,\dots,d\}} \left\{ \frac{\lambda_j - \lambda_{j-1}}{4\lambda_j - \lambda_{j-1}} \right\}. \tag{4.9}$$

LEMMA 4.3. *For all* $j = 1, \dots, d$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \#\mathcal{A}_n^j \le (\lambda_j - \lambda_{j-1}) H^j(A, \lceil Lq \rceil).$$

Proof. Take $j \in \{1, \dots, d\}$. For each $\tau = (\tau_{\lceil\lambda_{j-1}n\rceil+1}, \dots, \tau_{\lceil\lambda_j n\rceil}) \in \mathcal{A}_{n_r}^j$, choose $\kappa^\tau = (\omega^\tau_{\lceil\lambda_{j-1}n\rceil+1}, \dots, \omega^\tau_{\lceil\lambda_j n\rceil}) \in \mathcal{D}^{\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil}$ so that

$$(\eta_j(\omega^\tau_{\lceil\lambda_{j-1}n\rceil+1}), \dots, \eta_j(\omega^\tau_{\lceil\lambda_j n\rceil})) = \tau$$

and

$$\Omega(A, N, q) \cap \sigma^{-\lceil\lambda_{j-1}n\rceil}[\omega^\tau_{\lceil\lambda_{j-1}n\rceil+1} \cdots \omega^\tau_{\lceil\lambda_j n\rceil}] \ne \emptyset. \tag{4.10}$$

We now let $\nu_n \in \mathcal{B}_{\sigma^{\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil}}(\Sigma)$ be the unique $(\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil)$th level Bernoulli measure satisfying

$$\nu_n([\omega_{\lceil\lambda_{j-1}n\rceil+1} \cdots \omega_{\lceil\lambda_j n\rceil}]) := \begin{cases} \dfrac{1}{\#\mathcal{A}_n^j} & \text{if } (\omega_{\lceil\lambda_{j-1}n\rceil+1}, \dots, \omega_{\lceil\lambda_j n\rceil}) = \kappa^\tau \text{ for some } \tau \in \mathcal{A}_n^j, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$\nu_n([\tau_{\lceil\lambda_{j-1}n\rceil+1} \cdots \tau_{\lceil\lambda_j n\rceil}]) := \begin{cases} \dfrac{1}{\#\mathcal{A}_n^j} & \text{for } (\tau_{\lceil\lambda_{j-1}n\rceil+1}, \dots, \tau_{\lceil\lambda_j n\rceil}) = \tau \in \mathcal{A}_n^j, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $h(\nu_n \circ \chi_j^{-1}, \sigma_j^{\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil}) = \log \#\mathcal{A}_n^j$ (see [**Wa**, Theorem 4.26]). Let

$$\mu_n := A_{\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil}(\nu_n).$$

By Lemma 4.2(i), each μ_n is ergodic, and by Lemma 4.2(ii), we have

$$\frac{1}{n} \log \#\mathcal{B}_n = \frac{\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil}{n} h(\mu_n^j \circ \chi_j^{-1}, \sigma_j). \tag{4.11}$$

By the definition of $\Omega(A, N, q)$ for all $\omega \in \Omega(A, N, q)$ and $n > N$, there exists α_n such that for all $l \le q$, we have

$$\left| \sum_{r=0}^{n-1} \varphi_l(\sigma^r(\omega)) - n \int \varphi_l \, d\alpha_n \right| < \frac{n}{q} \tag{4.12}$$

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

and hence for all $n > \lambda_1^{-1}N$, we have

$$\left|\sum_{r=\lceil\lambda_{j-1}n\rceil}^{\lceil\lambda_j n\rceil-1}\varphi_l(\sigma^r(\omega)) - (\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil)\int\varphi_l\,d\rho_n\right| < \frac{2\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil}{q}.$$

Thus, by equation (4.10), for all $\tau \in A_n^j$, all $\omega \in [\omega^\tau_{\lceil\lambda_{j-1}n\rceil+1}\cdots\omega^\tau_{\lceil\lambda_j n\rceil}]$ and all $l \le q$,

$$\left|\sum_{r=0}^{\lceil\lambda_j n\rceil-\lceil\lambda_{j-1}n\rceil-1}\varphi_l(\sigma^r(\omega)) - (\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil)\int\varphi_l\,d\rho_n\right| < \frac{2\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil}{q} + \sum_{r=0}^{\lceil\lambda_j n\rceil-\lceil\lambda_{j-1}n\rceil-1}\mathrm{var}_r(\varphi_l). \tag{4.13}$$

Since each φ_l is continuous, it follows that there exists some $M > \lambda_1^{-1}N$ such that for all $n \ge M$, all $\tau \in A_n^j$, all $\omega \in [\omega^\tau_{\lceil\lambda_{j-1}n\rceil+1}\cdots\omega^\tau_{\lceil\lambda_j n\rceil}]$ and all $l \le q$,

$$\left|\frac{1}{\lceil\lambda_j n\rceil - \lceil\lambda_{j-1}n\rceil}\sum_{r=0}^{\lceil\lambda_j n\rceil-\lceil\lambda_{j-1}n\rceil-1}\varphi_l(\sigma^r(\omega)) - \int\varphi_l\,d\rho_n\right| < \frac{3\lambda_j - \lambda_{j-1}}{q(\lambda_j - \lambda_{j-1})} < \frac{1}{\lceil Lq\rceil}. \tag{4.14}$$

Now, since ν_n is supported on sets of the form $[\omega^\tau_{\lceil\lambda_{j-1}n\rceil+1}\cdots\omega^\tau_{\lceil\lambda_j n\rceil}]$, with $\tau \in A_n^j$, it follows that for all $n \ge M$ and all $l \le q$,

$$\left|\int A_{\lceil\lambda_j n\rceil-\lceil\lambda_{j-1}n\rceil}(\varphi_l)\,d\nu_n - \int\varphi_l\,d\rho_n\right| < \frac{1}{\lceil Lq\rceil}.$$

Thus, by Lemma 4.2(iii), we have

$$\left|\int\varphi_l\,d\mu_n - \int\varphi_l\,d\rho_n\right| < \frac{1}{\lceil Lq\rceil}$$

for all $n \ge M$ and $l \le \lceil Lq\rceil$, $\mu_n \in U(A, \lceil Lq\rceil)$, and hence

$$h(\mu_n^j \circ \chi_j^{-1}, \sigma_j) \le H^j(A, \lceil Lq\rceil). \tag{4.15}$$

By equation (4.11), this proves the lemma. □

LEMMA 4.4. *For each $q \in \mathbb{N}$,*

$$\dim_{\mathcal{P}}\Pi(\Omega(A)) \le \frac{H^1(\sigma, A, \lceil Lq\rceil)}{\log a_1} + \sum_{j=2}^{d}\left(\frac{1}{\log a_j} - \frac{1}{\log a_{j-1}}\right)H^j(\sigma, A, \lceil Lq\rceil).$$

Proof. Combining Lemma 4.1 with Lemma 4.3, we have

$$\overline{\dim}_B\Pi(\Omega(A, N, q)) \le \frac{H^1(\sigma, A, \lceil Lq\rceil)}{\log a_1} + \sum_{j=2}^{d}\left(\frac{1}{\log a_j} - \frac{1}{\log a_{j-1}}\right)H^j(\sigma, A, \lceil Lq\rceil)$$

for each $N \in \mathbb{N}$. Moreover, since

$$\Pi(\Omega(A)) \subseteq \bigcup_{N\in\mathbb{N}}\Pi(\Omega(A, N, q)),$$

we may apply Proposition 4.1 to prove the lemma. □

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

LEMMA 4.5. *For each* $j = 1, \ldots, d$, $\lim_{l\to\infty} H^j(A, l) = H^j(\sigma, A)$.

Proof. Fix $j \in \{1, \ldots, d\}$. Clearly, $H^j(\sigma, A) \leq \liminf_{l\to\infty} H^j(A, l)$. Now, for each $l \in \mathbb{N}$, choose $\mu_l \in U(A, l)$ with $h_{\mu_l\circ\chi_j^{-1}}(\sigma_j) > H^j(A, l) - (1/l)$. Since $\mathcal{M}_\sigma(\Sigma)$ is compact, we may take a weak $*$ limit $\mu_\infty \in \mathcal{M}_\sigma(\Sigma)$. It follows from the fact that A is closed and $\mu_l \in U(A, l)$ for each $l \in \mathbb{N}$ that $\mu_\infty \in A$. Moreover, since entropy is upper semi-continuous (see [**Wa**, Theorem 8.2]),

$$\begin{aligned} H^j(\sigma, A) &\geq h_{\mu_\infty\circ\chi_j^{-1}}(\sigma_j) \\ &\geq \limsup_{l\to\infty} h_{\mu_l\circ\chi_j^{-1}}(\sigma_j) \\ &\geq \limsup_{l\to\infty} H^j(A, l). \end{aligned} \tag{4.16}$$

□

To complete the proof, we let $q \to \infty$ in Lemma 4.4. Applying Lemma 4.5, we have

$$\dim_{\mathcal{P}}\Pi(\Omega(A)) \leq \frac{H^1(\sigma, A)}{\log a_1} + \sum_{j=2}^{d}\left(\frac{1}{\log a_j} - \frac{1}{\log a_{j-1}}\right)H^j(\sigma, A). \tag{4.17}$$

This completes the proof of Theorem 7 and hence Theorems 5 and 6.

5. *The shape of the spectrum*

We now deduce several features of the shape of the packing spectrum.

COROLLARY 1. *Let* $\varphi : \Sigma \to \mathbb{R}$ *be a continuous real-valued potential which is not cohomologous to a constant. Then, the packing spectrum* $\alpha \mapsto \dim_{\mathcal{P}}\Pi(J_\varphi(\alpha))$ *is concave and continuous on the interval* $A(\varphi) = [\alpha_{\min}, \alpha_{\max}]$.

Proof. By Theorem 7, it suffices to show that for each $j = \{1, \ldots, d\}$, $H^j(\sigma, \varphi, \alpha)$ is concave and continuous. So, fix $j \in \{1, \ldots, d\}$. It follows from Lemma 4.5 that $H^j(\sigma, \varphi, \alpha)$ is upper semi-continuous. Moreover, $H^j(\sigma, \varphi, \alpha)$ is concave. Indeed, given α_-, $\alpha_+ \in A(\varphi)$ and $\delta > 0$, we may choose $\mu_-, \mu_+ \in \mathcal{M}_\sigma(\Sigma)$ such that $\int \varphi\, d\mu_- = \alpha_-$ and $\int \varphi\, d\mu_+ = \alpha_+$ $h(\mu_- \circ \chi_j^{-1}, \sigma_j) > H^j(\sigma, \varphi, \alpha_-) - \delta$, $h(\mu_+ \circ \chi_j^{-1}, \sigma_j) > H^j(\sigma, \varphi, \alpha_+) - \delta$. For each $t \in (0, 1)$, we let $\mu_t := (1-t)\mu_- + t\mu_+$ so that $\int \varphi\, d\mu_t = (1-t)\alpha_- + t\alpha_+$. Moreover, since the entropy map is affine (see [**Wa**, Theorem 8.1]),

$$\begin{aligned} H^j(\sigma, \varphi, (1-t)\alpha_- + t\alpha_+) &\geq h(\mu_t \circ \chi_j^{-1}, \sigma_j) \\ &= (1-t)h(\mu_- \circ \chi_j^{-1}, \sigma_j) + th(\mu_+ \circ \chi_j^{-1}, \sigma_j) \\ &\geq (1-t)H^j(\sigma_j, \varphi, \alpha_-) + tH^j(\sigma_j, \varphi, \alpha_+) - \delta. \end{aligned}$$

Letting $\delta \to 0$, we see that $\alpha \mapsto H^j(\sigma, \varphi, \alpha)$ is concave and hence lower semi-continuous. □

The following corollary gives a sufficient condition on φ for the packing spectrum to be analytic.

COROLLARY 2. *Suppose there is some Hölder continuous potential* $\tilde{\varphi} : \Sigma_d \to \mathbb{R}$ *such that* $\varphi = \tilde{\varphi} \circ \chi_d$. *Then,* $\alpha \mapsto \dim_{\mathcal{P}} E_\varphi(\alpha)$ *is strictly concave and real analytic on the interval* $A(\varphi)$.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Proof. Note that for each $j = 1, \ldots, d$, the projection $\chi_d \circ \chi_j^{-1} : \Sigma_j \to \Sigma_d$ is a well-defined Lipschitz function. Hence the real-valued potential $\varphi_j : \Sigma_j \to \Sigma_d$, given by $\varphi_j := \tilde{\varphi} \circ \chi_d \circ \chi_j^{-1}$ is Hölder continuous. It follows straightforwardly from $\varphi = \tilde{\varphi} \circ \chi_d$ that for each $j = 1, \ldots, d$,

$$H^j(\sigma, \varphi, \alpha) = \sup\left\{h_\mu(\sigma_j) : \mu \in \mathcal{M}_{\sigma_j}(\Sigma_j), \int \varphi_j \, d\mu = \alpha\right\}. \tag{5.1}$$

One can deduce from standard results that the right-hand side of (5.1) is strictly concave and analytic. Since φ is not cohomologous to a constant, and $\chi_j \circ \sigma = \sigma_j \circ \chi_j$, it is clear that no φ_j is cohomologous to a constant. Now fix $j \in \{1, \ldots, d\}$. By [**Bo**, Theorem 1.28], it follows that for each j, the Gibbs measure corresponding to φ_j is not the measure of maximal entropy on Σ_j. Now, for each $\alpha \in A(\varphi)$, consider the set

$$J_j(\alpha) := \left\{\omega \in \Sigma_j : \lim_{n\to\infty} \frac{1}{n} \sum_{r=0}^{n} \varphi_j(\sigma^r \omega) = \alpha\right\}, \tag{5.2}$$

where Σ_j is given the usual symbolic metric (see [**Bo**, Ch. 1]). By [**BSa**, Theorem 6], $\dim_{\mathcal{H}} J_j(\alpha)$ is equal to a constant multiple of the quantity on the right-hand side of (5.1). Since the Gibbs measure corresponding to φ_j is not the measure of maximal entropy, it follows from [**PW1**, Theorem 1] that $\alpha \mapsto \dim_{\mathcal{H}} J_j(\alpha)$ is strictly concave and real analytic on $(\alpha_{\min}, \alpha_{\max})$. Thus, the spectrum is strictly convex and real analytic on $(\alpha_{\min}, \alpha_{\max})$. By Corollary 1, the spectrum is continuous on $[\alpha_{\min}, \alpha_{\max}]$ and hence these properties extend to the full interval $[\alpha_{\min}, \alpha_{\max}]$. □

For each $j = 1, \ldots, d$, we let $\mathbf{b}_j$ denote the measure of maximal entropy on Σ_j. We conclude this section with a necessary and sufficient condition for the packing spectrum to attain the full packing dimension of the repeller. The proof is immediate from Theorem 7.

COROLLARY 3. *There exists some $\alpha \in A(\varphi)$ satisfying $\dim_{\mathcal{P}} E_\varphi(\alpha) = \Lambda$ if and only if $\int \varphi \, d\mu_1 = \int \varphi \, d\mu_2 = \cdots = \int \varphi \, d\mu_d$ for some $\mu_1, \ldots, \mu_d \in \mathcal{M}_\sigma(\Sigma)$ such that $\mu_j \circ \chi_j^{-1} = \mathbf{b}_j$ for each $j = 1, \ldots, d$.*

6. *Examples*

In this section, we consider two simple examples exhibiting interesting features of the packing spectrum.

As noted in the introduction, the packing and Hausdorff spectra need not coincide. This raises the question of whether there are any real-valued potentials $\varphi : \Sigma \to \mathbb{R}$ supported on Bedford–McMullen repellers for which $\dim_{\mathcal{H}}(\Lambda) < \dim_{\mathcal{P}}(\Lambda)$, and yet the Hausdorff and packing spectra for φ coincide. Our first example shows that this can indeed be the case. One consequence of this is that $\dim_{\mathcal{P}} E_\varphi(\alpha) = \dim_{\mathcal{H}} E_\varphi(\alpha) \le \dim_{\mathcal{H}} \Lambda < \dim_{\mathcal{P}} \Lambda$ for all $\alpha \in [\alpha_{\min}, \alpha_{\max}]$. So, the packing spectrum need not attain the full packing dimension of the repeller at any point. This is in contrast to the situation for the Hausdorff dimension where there is always some $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ for which $\dim_{\mathcal{H}}(E_\varphi(\alpha)) = \dim_{\mathcal{H}}(\Lambda)$, namely $\alpha = \int \varphi \, d\mu_*$, where μ_* is an invariant measure of full dimension (see [**BM1, Be, McM**]).

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Example 1. Take $a_1 = 3$, $a_2 = 2$ and $\mathcal{D} = \{(0, 0), (1, 1), (2, 0)\}$ and $\varphi : \Lambda \to \mathbb{R}$, defined by

$$\varphi(\omega) = \begin{cases} 1 & \text{if } \omega_1 = (1, 1), \\ 0 & \text{if } \omega_1 \neq (1, 1). \end{cases}$$

Then, $\dim_{\mathcal{H}}(\Lambda) < \dim_{\mathcal{P}}(\Lambda)$. However, for all $\alpha \in [0, 1]$,

$$\dim_{\mathcal{P}} E_{\varphi\circ\Pi}(\alpha) = \dim_{\mathcal{H}} E_{\varphi\circ\Pi}(\alpha) = \frac{-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)}{\log 2} + (1-\alpha)\frac{\log 2}{\log 3}.$$

Proof. $\dim_{\mathcal{H}}(\Lambda) < \dim_{\mathcal{P}}(\Lambda)$ follows from Theorem 2. By considering the $((1-\alpha)/2, \alpha, (1-\alpha)/2)$-Bernoulli measure, it follows from Theorem 4 that

$$\dim_{\mathcal{H}} E_{\varphi}(\alpha) \geq \frac{-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)}{\log 2} + (1-\alpha)\frac{\log 2}{\log 3}.$$

It is easy to see that

$$\begin{aligned} &\sup\left\{\sum_{i=1}^{3} -p_i \log p_i : p_i \in [0, 1], \sum_{i=1}^{3} p_i = 1, \, p_1 = \alpha\right\} \\ &\quad = -\alpha \log \alpha - (1-\alpha)\log\left(\frac{1-\alpha}{2}\right) \\ &\sup\left\{\sum_{i=1}^{2} -p_i \log p_i : p_i \in [0, 1], \sum_{i=1}^{2} p_i = 1, \, p_1 = \alpha\right\} \\ &\quad = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha). \end{aligned}$$

Moreover, it follows from the fact that φ is locally constant (i.e. $\varphi(\omega') = \varphi(\omega)$ for all $\omega, \omega' \in \Sigma$ with $\omega'_1 = \omega_1$), together with the Kolmogorov–Sinai theorem, that the suprema

$$\sup\left\{h_{\mu}(\sigma) : \mu \in \mathcal{B}_{\sigma}(\Sigma), \int \varphi \, d\mu = \alpha\right\},$$
$$\sup\left\{h_{\mu\circ\chi_2^{-1}}(\sigma_v) : \mu \in \mathcal{B}_{\sigma}(\Sigma), \int \varphi \, d\mu = \alpha\right\}$$

are both attained by Bernoulli measures. Thus, applying Theorem 7, we have

$$\dim_{\mathcal{P}} E_{\varphi\circ\Pi}(\alpha) \leq \frac{-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)}{\log 2} + (1-\alpha)\frac{\log 2}{\log 3}.$$

□

For our next example, we have identical a_1, a_2 and $\mathcal{D}$, along with a potential φ which is prima facie very close to our previous one. Indeed, for $\alpha \geq \frac{1}{2}$, the spectra for the two examples coincide (see Figure 2). However, our next example has a point of non-analyticity at $\alpha = \frac{1}{2}$, and for $\alpha < \frac{1}{2}$ the two packing spectra are very different. In particular, the packing spectrum attains the packing dimension of Λ and so rises above the Hausdorff spectrum.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

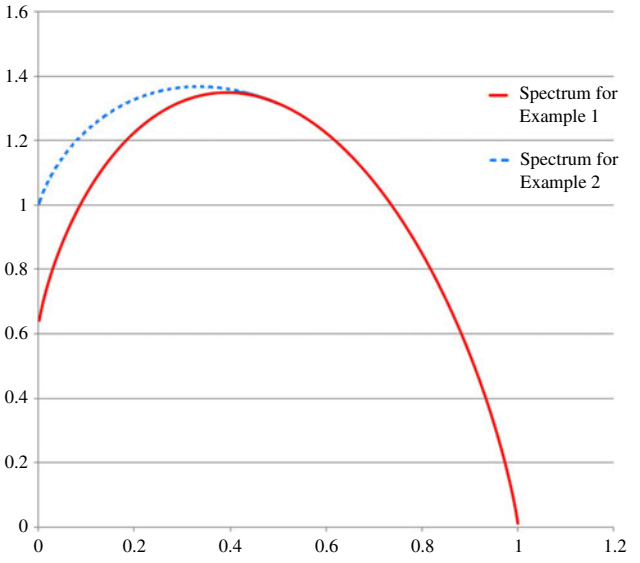


FIGURE 2. The packing spectra for Examples 1 and 2. The spectrum in Example 1 is given by the solid red line. Example 2 has a second-order phase-transition at $\alpha = \frac{1}{2}$. For $\alpha \leq \frac{1}{2}$, the spectrum in Example 2 is given by the broken blue line. For $\alpha > \frac{1}{2}$, the two spectra coincide and are both given by the solid red line.

Example 2. Take $a = 3$, $b = 2$ and $\mathcal{D} = \{(0, 0), (1, 1), (2, 0)\}$ and $\varphi : \Lambda \to \mathbb{R}$, defined by

$$\varphi(\omega) = \begin{cases} 1 & \text{if } \omega_1 = (2, 0), \\ 0 & \text{if } \omega_1 \neq (2, 0), \end{cases}$$

$$\dim_{\mathcal{P}} E_{\varphi \circ \Pi}(\alpha) = \begin{cases} \dfrac{-\alpha \log \alpha - (1-\alpha)\log(1-\alpha) - \alpha \log 2}{\log 3} + 1 & \text{for } \alpha \leq \dfrac{1}{2}, \\ \dfrac{-\alpha \log \alpha - (1-\alpha)\log(1-\alpha)}{\log 2} + (1-\alpha)\dfrac{\log 2}{\log 3} & \text{for } \alpha > \dfrac{1}{2}. \end{cases}$$

Moreover, $\alpha \mapsto \dim_{\mathcal{P}} E_{\varphi}(\alpha)$ is non-analytic and attains the full packing dimension $\dim_{\mathcal{P}}(\Lambda)$ at its maximum.

Proof. Note that

$$\sup\left\{\sum_{i=1}^{3} -p_i \log p_i : p_i \in [0, 1], \sum_{i=1}^{3} p_i = 1, p_1 = \alpha\right\}$$
$$= -\alpha \log \alpha - (1-\alpha)\log\left(\frac{1-\alpha}{2}\right),$$
$$\sup\left\{\sum_{i=1}^{2} -p_i \log p_i : p_i \in [0, 1], \sum_{i=1}^{2} p_i = 1, p_1 \geq \alpha\right\}$$
$$= \begin{cases} \log 2 & \text{for } \alpha \leq \dfrac{1}{2}, \\ -\alpha \log \alpha - (1-\alpha)\log(1-\alpha) & \text{for } \alpha > \dfrac{1}{2}. \end{cases}$$

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Moreover, since φ is locally constant, the following suprema are both attained by Bernoulli measures:

$$\sup\left\{h_\mu(\sigma) : \mu \in \mathcal{B}_\sigma(\Sigma), \int \varphi \, d\mu = \alpha\right\},$$
$$\sup\left\{h_{\mu\circ\chi_2^{-1}}(\sigma_v) : \mu \in \mathcal{B}_\sigma(\Sigma), \int \varphi \, d\mu = \alpha\right\}.$$

Thus, applying Theorem 7, we have

$$\dim_{\mathcal{P}} E_{\varphi\circ\Pi}(\alpha) = \begin{cases} \dfrac{-\alpha\log\alpha - (1-\alpha)\log(1-\alpha) - \alpha\log 2}{\log 3} + 1 & \text{for } \alpha \le \dfrac{1}{2}, \\ \dfrac{-\alpha\log\alpha - (1-\alpha)\log(1-\alpha)}{\log 2} + (1-\alpha)\dfrac{\log 2}{\log 3} & \text{for } \alpha > \dfrac{1}{2}. \end{cases}$$

Consequently, $\alpha \mapsto \dim_{\mathcal{P}} E_\varphi(\alpha)$ is non-analytic. This follows from the fact that the functions

$$\alpha \mapsto \frac{-\alpha\log\alpha - (1-\alpha)\log(1-\alpha) - \alpha\log 2}{\log 3} + 1,$$
$$\alpha \mapsto \frac{-\alpha\log\alpha - (1-\alpha)\log(1-\alpha)}{\log 2} + (1-\alpha)\frac{\log 2}{\log 3}$$

have distinct second derivatives at $\alpha = \frac{1}{2}$. It also follows from our expression for $\dim_{\mathcal{P}} E_\varphi(\alpha)$, together with Theorem 2, that the full packing dimension is attained at $\alpha = \frac{1}{3}$. □

7. *Generalizations and open questions*

In this section, we note some corollaries to Theorems 6 and 7. The first concerns sets of divergent points. Usually, one considers sets of points for which the Birkhoff average converges to a given value. However, given any non-empty closed convex subset $A \subseteq A(\varphi)$, one may consider the set $E_\varphi(A)$ of points $x \in \Lambda$ for which the set of accumulation points for the sequence $(A_n(\varphi)(x))_{n\in\mathbb{N}}$ is equal to A. In the conformal setting, both $\dim_{\mathcal{H}} E_\varphi(A)$ and $\dim_{\mathcal{P}} E_\varphi(A)$ have been well studied in a series of papers due to Olsen and Winter [**Ol1, Ol2, Ol3, OlWi**]. This follows work by Barreira and Schmeling [**BSch**] showing that, given finitely many continuous potentials $\varphi_1, \ldots, \varphi_N$ on a conformal repeller Λ, for which each $A(\varphi_i)$ consists of at least two points, the Hausdorff dimension (and hence packing dimension) of the set of all points for which none of the Birkhoff averages for $\varphi_1, \ldots, \varphi_N$ converge is of full Hausdorff dimension. We note that [**Ol3**, Theorem 4.3] implies that the set of points $x \in \Lambda$ for which the Birkhoff average $(A_n(\varphi)(x))_{n\in\mathbb{N}}$ does not converge, for any continuous potential φ for which $A(\varphi)$ consists of at least two points again, has full Hausdorff dimension. By a similar argument, along with some ideas from [**KP**], one can extend this result to self-affine Sierpiński sponges.

One application of Theorem 7 is to determine the packing dimension of the sets $E_\varphi(A)$ for self-affine Sierpiński sponges. We define, for $k = 1, \ldots, d$,

$$H^k(f, \varphi, A) := \sup\left\{h_{\mu\circ\pi_k^{-1}}(f_k) : \mu \in \mathcal{M}(f, \Lambda), \int \varphi \, d\mu \in A\right\}. \tag{7.1}$$

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

THEOREM 8. *Let Λ be a self-affine Sierpiński sponge. Let $\varphi : \Lambda \to \mathbb{R}^N$ be some continuous potential. Then, given any non-empty closed convex subset $A \subseteq A(\varphi)$, we have*

$$\dim_{\mathcal{P}} E_\varphi(A) = \frac{H^1(f, \varphi, A)}{\log a_1} + \sum_{k=2}^{d} \left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}} \right) H^k(f, \varphi, A).$$

In contrast, very little is known concerning the Hausdorff dimension of $E_\varphi(A)$ for self-affine Λ, aside from the special case where A is a singleton, and it would be very interesting to see if one could obtain a formula for $\dim_{\mathcal{H}} E_\varphi(A)$ for arbitrary non-empty closed convex subsets of $\mathcal{M}(\Lambda, f)$.

Theorem 7 also implies some results concerning the packing spectrum for the local dimension of a Bernoulli measure on a self-affine Sierpiński sponge. To each Bernoulli measure μ on Σ, we associate the corresponding probability vector $(p_{i_1 \cdots i_d})_{(i_1, \ldots, i_d) \in \mathcal{D}}$ in the usual way. Given $t \in \{1 \ \ldots, d\}$, we let $p_{i_1 \cdots i_t}$ denote the sum of all $p_{j_1 \cdots j_d}$ for which $(j_1, \ldots, j_t) = (i_1, \ldots, i_t)$.

For $j = 1, \ldots, d$, we define a potential $P_j : \Sigma \to \mathbb{R}$ by

$$P_j(\omega) := \begin{cases} \dfrac{\log p_{\chi_j(\omega_1)} / p_{\chi_{j+1}(\omega_1)}}{\log a_j} & \text{if } j \neq d, \\ \dfrac{\log p_{\chi_d(\omega_1)}}{\log a_d} & \text{if } j = d. \end{cases} \tag{7.2}$$

Clearly, $\mathrm{var}_1(P_j) = 0$ for each j and, as such, P_j is continuous. Let $P : \Sigma \to \mathbb{R}^d$ denote the potential $\omega \mapsto (P_j(\omega))_{j=1}^d$. We shall assume the very strong separation condition (see [**Ol4**, Condition (II)]).

LEMMA 7.1. *Suppose that μ is a Bernoulli measure on a self-affine Sierpiński sponge which satisfies the very strong separation condition. Then, for all $x = \Pi(\omega) \in \Lambda$, we have*

$$\lim_{r \to \infty} \frac{\log \mu(B(x, r))}{\log r} = \sum_{j=1}^{d} \frac{1}{\lceil \lambda_j n \rceil} \sum_{l=0}^{\lceil \lambda_j n \rceil - 1} P_j(\sigma^l \omega). \tag{7.3}$$

Proof. See [**Ol4**, Theorem 6.2.2]. □

Olsen [**Ol4**, Conjecture 4.1.7] conjectured that the packing spectrum for a Bernoulli measure on a self-affine Sierpiński sponge is given by the Legendre transform of a certain auxiliary function (see [**Ol4**, §3.1] for details). In particular, this conjecture would imply that the packing spectrum for the local dimension always peaks at the full packing dimension of the attractor Λ (see [**Ol4**, Theorem 3.3.2(ix)] and note that $\gamma(0) = \dim_{\mathcal{P}} \Lambda$ by Theorem 2). Theorem 7 provides us with the following counterexample.

Example 3. Take $a_1 = 4, a_2 = 3$ and $\mathcal{D} = \{(0, 0), (2, 2), (3, 0)\}$, and let ν be the Bernoulli measure obtained by taking $p_{00} = p_{30} = 1/4$ and $p_{22} = 1/2$. Let $\alpha_{\min} := \log 2 / \log 3$ and $\alpha_{\max} := \log 2 / \log 3 + 1/2$, and, for all $\alpha \in [\alpha_{\min}, \alpha_{\max}]$, we define $\rho(\alpha) := 2(\alpha - \log 2 / \log 3)$. Then, $\dim_{\mathcal{H}}(\Lambda) < \dim_{\mathcal{P}}(\Lambda)$. However, for all $\alpha \in [\alpha_{\min}, \alpha_{\max}]$,

$$\begin{aligned} \dim_{\mathcal{P}} D_\nu(\alpha) &= \dim_{\mathcal{H}} D_\nu(\alpha) \\ &= \frac{-\rho(\alpha) \log \rho(\alpha) - (1 - \rho(\alpha)) \log(1 - \rho(\alpha))}{\log 4} + \frac{1}{2}(1 - \rho(\alpha)). \end{aligned}$$

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Proof. Theorem 2 implies $\dim_{\mathcal{H}}(\Lambda) < \dim_{\mathcal{P}}(\Lambda)$. Applying Lemma 7.1, we see that for all $\alpha \in [\alpha_{\min}, \alpha_{\max}]$, $\Pi(\omega) \in D_\nu(\alpha)$ if and only if

$$\mathcal{V}(\omega) \subseteq \{\mu \in \mathcal{M}_\sigma(\Sigma) : \mu([(0,0)]) + \mu([(3,0)]) = \rho(\alpha)\}. \tag{7.4}$$

Now proceed as in Example 1. □

Theorem 7 also implies the following lower bound for the packing spectrum for the local dimension.

PROPOSITION 7.1. *Suppose that μ is a Bernoulli measure on a self-affine Sierpiński sponge which satisfies the very strong separation condition. Then,*

$$\dim_{\mathcal{P}} D_\mu(\alpha) \geq \sup\left\{\frac{H^1(\sigma, \underline{P}, \underline{\alpha})}{\log a_1} + \sum_{k=2}^{d}\left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}}\right)H^k(f, \underline{P}, \underline{\alpha})\right\},$$

where the supremum is taken over all $\underline{\alpha} = (\alpha_j)_{j=1}^d \in \mathbb{R}^d$ for which $\sum_{j=1}^d \alpha_j = \alpha$.

Proof. It follows from Lemma 7.1 that $E_{\underline{P}}(\underline{\alpha}) \subseteq D_\mu(\alpha)$ for each $\underline{\alpha} = (\alpha_j)_{j=1}^d \in \mathbb{R}^d$ with $\sum_{j=1}^d \alpha_j = \alpha$. Consequently, the result follows from Theorem 7. □

For a rather limited class of Bernoulli measures, we obtain an equality.

DEFINITION 7.1. We say that a Bernoulli measure μ on a self-affine Sierpiński sponge is one-dimensional if there exists some $k \in \{1, \ldots, d\}$ for which the probability vector $(p_{i_1 \cdots i_d})_{(i_1,\ldots,i_d)\in\mathcal{D}}$ associated to μ satisfies $p_{i_1 \cdots i_{d+1-q}} / p_{i_1 \cdots i_{d-q}} = \#\eta_{q+1}(\mathcal{D})/\#\eta_q(\mathcal{D})$ for all $(i_1, \ldots, i_{d+1-q}) \in \eta_{q+1}(\mathcal{D})$ and all $q \in \{1, \ldots, d-1\}\backslash\{k\}$ and each $p_i = 1/\#\eta_d(\mathcal{D})$ for $i \in \eta_d(\mathcal{D})$, provided $k \neq d$.

Now, if μ is a one-dimensional Bernoulli measure on a self-affine Sierpiński sponge, then for each $j \neq k$, P_j will be equal to an explicit constant c_j. Let $\tilde{P}$ denote the potential $P_k + \sum_{j\neq k} c_j$.

THEOREM 9. *Suppose that μ is a one-dimensional Bernoulli measure on a self-affine Sierpiński sponge which satisfies the very strong separation condition. Then,*

$$\dim_{\mathcal{P}} D_\mu(\alpha) = \frac{H^1(\sigma, \tilde{P}, \alpha)}{\log a_1} + \sum_{k=2}^{d}\left(\frac{1}{\log a_k} - \frac{1}{\log a_{k-1}}\right)H^k(f, \tilde{P}, \alpha).$$

Proof. It follows from Lemma 7.1 that $D_\mu(\alpha) = E_{\tilde{P}}(\alpha)$. Hence, the result follows from Theorem 7. □

We emphasize that the class of one-dimensional Bernoulli measures is really very limited and the techniques of this paper are insufficient for determining $\dim_{\mathcal{P}} D_\mu(\alpha)$ for more general classes of Bernoulli measures. The reason for this extra level of difficulty is that one is essentially dealing with a sum of Birkhoff averages taken at multiple timescales (see Lemma 7.1). It seems unlikely that the lower bound given in Proposition 7.1 is optimal. As such, it remains an open question to determine the packing spectrum for the local dimension on a self-affine Sierpiński sponge.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

Acknowledgements. The author would like to thank Thomas Jordan and Michał Rams for all their help in preparing this paper. This project was started in the Instytut Matematyczny PAN and I would also like to thank the Institute, especially Feliks Przytycki, for their kind hospitality. I would also like to thank the Engineering and Physical Sciences Research Council and the Conformal Structures and Dynamics network for their financial support. In addition, I would like to thank the referee, whose careful reading of this paper has greatly improved the presentation.

REFERENCES

[BOS] I. Baek, L. Olsen and N. Snigireva. Divergence points of self-similar measures and packing dimension. *Adv. Math.* **214**(1) (2007), 267–287.

[BF] J. Barral and D. Feng. Weighted thermodynamic formalism and applications. 2009, arXiv:0909.4247v1.

[BM1] J. Barral and M. Mensi. Multifractal analysis of Birkhoff averages on 'self-affine' symbolic spaces. *Nonlinearity* **21**(10) (2008), 2409–2425.

[BM2] J. Barral and M. Mensi. Gibbs measures on self-affine Sierpinski carpets and their singularity spectrum. *Ergod. Th. & Dynam. Sys.* **27**(5) (2007), 1419–1443.

[BSa] L. Barreira and B. Saussol. Variational principles and mixed multifractal spectra. *Trans. Amer. Math. Soc.* **353**(10) (2001), 3919–3944.

[BSch] L. Barreira and J. Schmeling. Sets of non-typical points have full topological entropy and full Hausdorff dimension. *Israel J. Math.* **116** (2000), 29–70.

[Be] T. Bedford. PhD Thesis: Crinkly curves, Markov partitions and box dimension of self-similar sets. *PhD Thesis*, University of Warwick, 1984.

[Bo] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, 2nd revised edn. *(Lecture Notes in Mathematics, 470)*. Springer, Berlin, 2008.

[Ed1] G. A. Edgar. *Measure, Topology, and Fractal Geometry*, 2nd edn. *(Undergraduate Texts in Mathematics)*. Springer, New York, 2008.

[Ed2] G. A. Edgar. *Integral, Probability, and Fractal Measures*. Springer, New York, 1998.

[Fa1] K. Falconer. The Hausdorff dimension of self-affine fractals. *Math. Proc. Cambridge Philos. Soc.* **103**(2) (1988), 339–350.

[Fa2] K. Falconer. *Fractal Geometry. Mathematical Foundations and Applications*, 2nd edn. John Wiley & Sons, Inc., Hoboken, NJ, 2003.

[Fa3] K. Falconer. *Techniques in Fractal Geometry*. John Wiley & Sons, Ltd., Chichester, 1997.

[FFW] A. Fan, D. Feng and J. Wu. Recurrence, dimension and entropy. *J. Lond. Math. Soc.* (2) **64**(1) (2001), 229–244.

[FLW] D. Feng, K. Lau and J. Wu. Ergodic limits on the conformal repellers. *Adv. Math.* **169**(1) (2002), 58–91.

[GR] K. Gelfert and M. Rams. The Lyapunov spectrum of some parabolic systems. *Ergod. Th. & Dynam. Sys.* **29**(3) (2009), 919–940.

[JJOP] A. Johansson, T. Jordan, A. Oberg and M. Pollicott. Multifractal analysis of non-uniformly hyperbolic systems. *Israel J. Math.* **177** (2010), 125–144.

[JR] T. Jordan and M. Rams. Multifractal analysis for Bedford–McMullen carpets. *Math. Proc. Cambridge Philos. Soc.* **150**(1) (2011), 147–156.

[JS] T. Jordan and K. Simon. Multifractal analysis of Birkhoff averages for some self-affine IFS. *Dyn. Syst.* **22**(4) (2007), 469–483.

[Ki] J. King. The singularity spectrum for general Sierpinski carpets. *Adv. Math.* **116**(1) (1995), 1–11.

[KP] R. Kenyon and Y. Peres. Measures of full dimension on affine-invariant sets. *Ergod. Th. & Dynam. Sys.* **16**(2) (1996), 307–323.

[Mat] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability (Cambridge Studies in Advanced Mathematics, 44)*. Cambridge University Press, Cambridge, 1995.

[McM] C. McMullen. The Hausdorff dimension of general Sierpinski carpets. *Nagoya Math. J.* **96** (1984), 1–9.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press

[Ni] O. Nielsen. The Hausdorff and packing dimensions of some sets related to Sierpiński carpets. *Canad. J. Math.* **51**(5) (1999), 1073–1088.

[Ol1] L. Olsen. Multifractal analysis of divergence points of deformed measure theoretical Birkhoff averages. *J. Math. Pures Appl.* (9) **82**(12) (2003), 1591–1649.

[Ol2] L. Olsen. Multifractal analysis of divergence points of deformed measure theoretical Birkhoff averages III. *Aequationes Math.* **71**(1–2) (2006), 29–53.

[Ol3] L. Olsen. Multifractal analysis of divergence points of deformed measure theoretical Birkhoff averages. IV: divergence points and packing dimension. *Bull. Sci. Math.* **132**(8) (2008), 650–678.

[Ol4] L. Olsen. Self-affine multifractal Sierpiński sponges in $\mathbb{R}^d$. *Pacific J. Math.* **183**(1) (1998), 143–199.

[OlWi] L. Olsen and S. Winter. Multifractal analysis of divergence points of deformed measure theoretical Birkhoff averages II: non-linearity, divergence points and Banach space valued spectra. *Bull. Sci. Math.* **131**(6) (2007), 518–558.

[Per] Y. Peres. The packing measure of self-affine carpets. *Math. Proc. Cambridge Philos. Soc.* **115**(3) (1994), 437–450.

[Pes] Y. Pesin. *Dimension Theory in Dynamical Systems. Contemporary Views and Applications (Contemporary Views and Applications. Chicago Lectures in Mathematics)*. University of Chicago Press, Chicago, IL, 1997.

[PW1] Y. Pesin and H. Weiss. The multifractal analysis of Birkhoff averages and large deviations. *Global Analysis of Dynamical Systems*. Institute of Physics, Bristol, 2001, pp. 419–431.

[PW2] Y. Pesin and H. Weiss. The multifractal analysis of Gibbs measures: motivation, mathematical foundation and examples. *Chaos* **7**(1) (1997), 89–106.

[Wa] P. Walters. *An Introduction to Ergodic Theory (Graduate Texts in Mathematics, 79)*. Springer, New York–Berlin, 1982.

https://doi.org/10.1017/S0143385711000368 Published online by Cambridge University Press