

## THE DIMENSION OF PROJECTIONS OF SELF-AFFINE SETS AND MEASURES

Kenneth Falconer and Tom Kempton

University of St Andrews, Mathematical Institute  
North Haugh, St Andrews, Fife, KY16 9SS, Scotland; kjf@st-andrews.ac.uk

University of St Andrews, Mathematical Institute  
North Haugh, St Andrews, Fife, KY16 9SS, Scotland; tmwk@st-andrews.ac.uk

**Abstract.** Let  $E$  be a plane self-affine set defined by affine transformations with linear parts given by matrices with positive entries. We show that if  $\mu$  is a Bernoulli measure on  $E$  with  $\dim_H \mu = \dim_L \mu$ , where  $\dim_H$  and  $\dim_L$  denote Hausdorff and Lyapunov dimensions, then the projection of  $\mu$  in all but at most one direction has Hausdorff dimension  $\min\{\dim_H \mu, 1\}$ . We transfer this result to sets and show that many self-affine sets have projections of dimension  $\min\{\dim_H E, 1\}$  in all but at most one direction.

### 1. Introduction

Marstrand's projection theorem states that, given a set  $E \subset \mathbf{R}^n$ , for almost every  $m$ -dimensional linear subspace  $K$  of  $\mathbf{R}^m$

$$(1.1) \quad \dim_H \pi_K E = \min\{\dim_H E, m\},$$

where  $\pi_K$  denotes orthogonal projection onto  $K$  and  $\dim_H$  denotes Hausdorff dimension, see [6, 17, 18]. The analogous projection theorem for measures, generally proved using potential theoretic methods, see [16], is that for a finite Borel measure  $\mu$  on  $\mathbf{R}^n$ ,

$$(1.2) \quad \dim_H \pi_K \mu = \min\{\dim_H \mu, m\},$$

for almost all subspaces  $K$ , where the (*lower*) Hausdorff dimension of a probability measure is given in terms of dimensions of sets by

$$\dim_H \nu = \inf\{\dim_H U : \nu(U) = 1\}.$$

However, these projection theorems are no help in identifying the subspaces  $K$ , if any, for which (1.1) or (1.2) fail. Recently there has been a great deal of interest in finding the 'exceptional directions' for projections of fractals and fractal measures of various types, especially those with a recursive structure or which are dynamically generated, see the survey [7] and references therein. In particular, various classes of measures and sets have been shown to have every projection satisfying (1.1). For example, for self-similar sets satisfying the open set condition, (1.1) holds for all subspaces  $K$  provided that the group generated by the orthonormal components of the defining similarities is dense in the orthogonal group  $O(n)$ , with similar results for measures, see [8, 19, 13]. On the other hand, if the group is not dense then there are always some exceptional directions, see [10].

It is natural to ask under what circumstances (1.1) holds for self-affine sets and (1.2) for self-affine measures (by which we mean Bernoulli measures on self-affine

---

<https://doi.org/10.5186/aasfm.2017.4232>

2010 Mathematics Subject Classification: Primary 28A30.

Key words: Dimension, projection, self-affine set, Furstenberg measure,  $r$ -scale entropy.

This research was supported by EPSRC grant EP/K029061/1.

sets) in all or virtually all directions. Dimensional analysis in the self-affine case is more awkward than in the self-similar case, not least because the dimensions of the sets or measures themselves are not continuous in the defining affine transformations. Nevertheless, the dimensions of many specific self-affine sets and also of generic self-affine sets in certain families are now known. In particular, the Hausdorff dimension  $\dim_H E$  of a self-affine set  $E$  ‘often’ equals its affinity dimension  $\dim_A E$ , defined in (2.2)–(2.3) below, see [5].

Much of the work to date on projections of self-affine sets concerns projections of plane self-affine carpets, that is where the defining affine maps preserve both horizontal and vertical axes; see [1, 11] for various examples. Perhaps not surprisingly, for many carpets (1.1) fails only for projections in the horizontal and/or vertical directions.

In this paper we consider projections of plane self-affine sets and measures where the linear parts of the self-affine maps may be represented by matrices with all entries positive, in other words where the linear parts map the positive quadrant strictly into itself. We show that in many cases projections onto lines in *all* directions (in some cases all but one direction) satisfy (1.1) or (1.2). Many self-affine sets and measures fit into this context, both in the ‘generic setting’ (provided that the affine maps are sufficiently contracting), see [4, 22], and for many specific cases including those presented in [2, 9, 12].

To study projections of a set one normally examines the projections of a suitable measure on the set. Thus, we first study projections of a Bernoulli measure  $\mu$  supported by a self-affine set  $E$ . Then  $\mu$  induces a measure  $\mu_F$ , known as a Furstenberg measure, on the space of line directions. Recent results on the dynamical structure of self-affine sets [2, 9] imply that the projections of  $\mu$  are exact dimensional in  $\mu_F$ -almost all directions. Together with lower bounds for the dimension of projections from [13] involving  $r$ -scale entropies we conclude that the dimensions of projections of measures are well-behaved in all but perhaps one direction. The results on projections of self-affine sets follow by supporting suitable Bernoulli measures on  $E$ .

## 2. Preliminaries

In this section we introduce the self-affine sets that we consider along with pertinent notation. Let  $A_1, \dots, A_k$  be a collection of  $2 \times 2$  matrices of Euclidean norm less than 1 with strictly positive entries, and let  $d_1, \dots, d_k \in \mathbf{R}^2$ . Then the maps  $T_i: \mathbf{R}^2 \rightarrow \mathbf{R}^2$  ( $i = 1, \dots, k$ ) given by

$$(2.1) \quad T_i(x) = A_i(x) + d_i \quad (x \in \mathbf{R}^2)$$

are affine contractions which form an iterated function system (IFS). By standard IFS theory, see [6, 14] there exists a unique non-empty compact set  $E$  satisfying

$$E = \bigcup_{i=1}^k T_i(E),$$

termed a *self-affine set*.

There is a natural coding on  $E$  and its usual iterated construction. Let  $\Lambda := \{1, \dots, k\}$ , let  $\Sigma^* = \bigcup_{n=0}^{\infty} \Lambda^n$  be the space of finite words formed by elements of  $\Lambda$ , and let  $\Sigma = \Lambda^{\mathbf{N}}$  be the corresponding space of infinite words. Associated with each  $a_1 \cdots a_n \in \Sigma^*$  is the *cylinder*

$$[a_1 \cdots a_n] = \{a_1 \cdots a_n b_{n+1} b_{n+2}, \dots : b_i \in \Lambda\} \subset \Sigma;$$

the cylinders form a basis of open and closed sets for the natural topology on  $\Sigma$ .

We abbreviate compositions of the  $T_i$  by  $T_{a_1 \dots a_n} := T_{a_1} \circ \dots \circ T_{a_n}$  for  $a_1 \dots a_n \in \Sigma^*$ . Let  $D \subset \mathbf{R}^2$  be the unit disk so that

$$D_{a_1 \dots a_n} := T_{a_1 \dots a_n}(D)$$

are the ellipses obtained as images of  $D$  under compositions of the  $T_i$ . By scaling we may assume that  $T_i(D) \subset D$  for all  $i$ , in which case the self-affine set  $E$  may be written as

$$E = \bigcap_{n=0}^{\infty} \bigcup_{a_1 \dots a_n} T_{a_1 \dots a_n}(D).$$

Let  $\mathbf{PR}^1 = (\mathbf{R}^2 \setminus 0) / \sim$  be 1-dimensional real projective space where  $\sim$  identifies points on each line through the origin. We may parameterise  $\mathbf{PR}^1$  by the angle between lines and the horizontal axis, and these angles define a metric  $d$  on  $\mathbf{PR}^1$  in the obvious way. Each matrix  $A_i^{-1}$  induces a projective linear transformation  $\phi_i: \mathbf{PR}^1 \rightarrow \mathbf{PR}^1$  given by  $\phi_i := A_i^{-1} / \sim$ , that is,  $A_i^{-1}$  maps straight lines at angle  $\theta$  to straight lines at angle  $\phi_i(\theta)$ . Since the matrices  $A_i$  are strictly positive, each  $\phi_i$  restricted to the negative quadrant  $\mathcal{Q}_2$  is a contraction so that the system of contractions  $(\phi_i)_{i=1}^k$  is an iterated function system on  $\mathcal{Q}_2$ . Let  $F$  be the attractor of this IFS, that is the non-empty compact subset of  $\mathcal{Q}_2$  such that  $F = \bigcup_{i=1}^k T_i(F)$ .

Now let  $\mu$  be a Bernoulli probability measure on  $\Sigma$ ; for notational convenience we also write  $\mu$  for the corresponding measure on  $E$ , that is its image under the map  $(a_1 a_2 \dots) \mapsto \bigcap_{n=0}^{\infty} T_{a_1 \dots a_n}(D)$ . The *Furstenberg measure*  $\mu_F$  is defined to be the stationary probability measure supported by  $K \subset \mathbf{RP}^1$  associated to the maps  $\phi_i$  chosen according to the measure  $\mu$ , see for example, [3]. Specifically,

$$\mu_F(U) = \sum_{i=1}^k \mu[i] \mu_F(\phi_i^{-1}(U)),$$

for all Borel sets  $U$ . The Furstenberg measure is central to the proofs of our results.

Ideally we would like to be able to express the Hausdorff dimension of a self-affine set  $E$  in terms of its defining IFS maps (2.1), as can be done for self-similar sets subject to a separation condition. (For the definition and basic properties of Hausdorff dimension, which we denote by  $\dim_H$ , see for example [6, 18].) A natural candidate for  $\dim_H E$  is the affinity dimension defined in [4] in terms of the linear parts of the IFS maps. Let  $\alpha_1(A) \geq \alpha_2(A) \geq 0$  be the *singular values* of a linear mapping or matrix  $A$  on  $\mathbf{R}^2$ , that is the lengths of the major and minor semiaxes of the ellipse  $A(D)$  or equivalently the positive square roots of the eigenvalues of  $AA^T$ . For  $0 \leq s \leq 2$  we define the *singular value function* of  $A$  by

$$\phi^s(A) = \begin{cases} \alpha_1^s, & 0 < s \leq 1, \\ \alpha_1 \alpha_2^{s-1}, & 1 \leq s \leq 2, \\ (\det A)^{s/2}, & 2 \leq s. \end{cases}$$

The submultiplicativity of the singular value functions enables us to define the *affinity dimension* of a set  $E \subset \mathbf{R}^2$  defined by the IFS (2.1) by

$$(2.2) \quad \dim_A(A_1, \dots, A_k) \equiv \dim_A E = \left\{ s: \lim_{n \rightarrow \infty} \left( \sum_{a_1 \dots a_n \in \Lambda^n} \phi^s(A_{a_1 \dots a_n}) \right)^{1/n} = 1 \right\}$$

or equivalently

$$(2.3) \quad \dim_A(A_1, \dots, A_k) \equiv \dim_A E = \inf \left\{ s : \sum_{n=1}^{\infty} \sum_{a_1 \dots a_n \in \Lambda^n} \phi^s(A_{a_1 \dots a_n}) < \infty \right\}.$$

Note that the affinity dimension depends on the  $\{A_i\}$  that occur in the iterated function system that defines  $E$ , rather than on  $E$  itself, but referring to  $\dim_A E$  generally causes little problem. For every self-affine set  $E$  it is the case that  $\dim_H E \leq \dim_A E$ , but equality holds in many cases, both generic and specific, see the survey [5] and recent examples in [2, 9].

We will also refer to Lyapunov dimension, which reflects the geometry of the self-affine set  $E$  relative to the measure  $\mu$ . The *Lyapunov exponents*  $\lambda_1(\mu), \lambda_2(\mu)$  are constants such that, for  $\mu$ -almost every  $a_1 a_2 \dots \in \Sigma$ ,

$$(2.4) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log \alpha_i(a_1 \dots a_n) = \lambda_i \quad (i = 1, 2).$$

The *Lyapunov dimension* of  $\mu$  is given by

$$(2.5) \quad \dim_L \mu := \begin{cases} \frac{h_\sigma(\mu)}{-\lambda_1(\mu)}, & h_\sigma(\mu) \leq -\lambda_1(\mu), \\ 1 + \frac{h_\sigma(\mu) + \lambda_1(\mu)}{-\lambda_2(\mu)}, & h_\sigma(\mu) \geq -\lambda_1(\mu), \end{cases}$$

where  $h_\sigma(\mu)$  is the Kolmogorov–Sinai entropy of the system  $(\Sigma, \sigma, \mu)$  with  $\sigma$  the left shift on  $\Sigma$ . Note that  $\dim_L \mu$  depends only on the matrices  $\{A_1, \dots, A_k\}$  and the measure  $\mu$ . It is always the case that  $\dim_H \mu \leq \dim_L \mu$ , see [15].

To obtain our dimension estimates, we will use *r-scale entropies*, given by

$$H_r(\nu) := - \int_{[-1,1]} \log \nu(B_r(x)) \, d\nu(x),$$

where  $\nu$  is a probability measure on  $[-1, 1]$  and  $r > 0$ . A measure  $\nu$  is termed *exact dimensional of dimension  $\beta$*  if

$$(2.6) \quad \lim_{r \rightarrow 0} \frac{\log \nu(B_r(x))}{\log r} = \beta$$

for  $\nu$ -almost all  $x$ . In particular, if  $\nu$  is exact dimensional, then

$$(2.7) \quad \dim_H \nu = \beta = \lim_{r \rightarrow 0} \frac{H_r(\nu)}{-\log r}.$$

### 3. Statement of results

We define  $B \subset \mathbf{PR}^1$  by  $B = \mathbf{PR}^1 \setminus \{\theta_0\}$  if all matrices  $A_i$  have a common maximal eigenvector in some direction  $\theta_0$ , and by  $B = \mathbf{PR}^1$  otherwise. The set  $\mathbf{PR}^1 \setminus B$  consists of at most one angle and represents the set of possible exceptional directions for our projection theorems.

It was proved in [2, Proposition 3.3] that there exists a constant  $\beta(\mu)$  such that for  $\mu_F$ -almost every  $\theta$  the projected measure  $\pi_\theta(\mu)$  is exact dimensional with dimension  $\beta(\mu)$ . This allows us to state our main theorem.

**Theorem 3.1.** *Let  $\mu$  be a Bernoulli measure on a planar self-affine set defined by an IFS (2.1) with strictly positive matrices  $A_i$ . Then for all  $\theta \in B$*

$$\dim_H \pi_\theta(\mu) \geq \beta(\mu).$$

This yields the following corollaries.

**Corollary 3.2.** *Let  $\mu$  be a Bernoulli measure on a planar self-affine set associated to strictly positive matrices  $A_i$ . Suppose  $\dim_H \mu = \dim_L \mu$ . Then for all  $\theta \in B$ ,*

$$\dim_H \pi_\theta(\mu) = \min\{\dim_H \mu, 1\}.$$

*Proof.* If we can show that

$$(3.1) \quad \dim_H \mu = \dim_L \mu \implies \beta(\mu) = \min\{\dim_H \mu, 1\}.$$

the conclusion follows directly from Theorem 3.1. Under the additional assumption that the self-affine set satisfies the strong separation condition, (3.1) along with the converse implication was established in [9]. However, the strong separation condition is not required for the implication (3.1). This is because the left hand side of inequality (4.2) of [9, Lemma 4.2] holds even without strong separation.  $\square$

Theorem 3.1 can be applied to the projection of self-affine sets. The next corollary shows that if the set supports Bernoulli measures with dimensions approximating the affinity dimension there is no dimension drop for projections of  $E$  except possibly in one direction.

**Corollary 3.3.** *Let  $E$  be a planar self-affine set defined by an IFS (2.1) with strictly positive matrices  $A_i$ . If there exists a sequence  $(\mu_n)$  of Bernoulli measures supported on  $E$  with  $\dim_H \mu_n \rightarrow \dim_A E$  then, for all  $\theta \in B$ ,*

$$(3.2) \quad \dim_H \pi_\theta(E) = \min\{\dim_H E, 1\}.$$

*Proof.* Let  $\mu_{n,F}$  denote the Furstenberg measure associated to  $\mu_n$ , and let  $\beta_n$  denote the value of  $\dim_H \pi_\theta(\mu_n)$  that occurs for  $\mu_{n,F}$ -almost every  $\theta$  by [2, Proposition 3.3]. Then Theorem 3.1 implies that for all  $\theta \in B$ ,

$$\dim_H \pi_\theta(E) \geq \lim_{n \rightarrow \infty} \beta_n = \min\{\dim_A E, 1\},$$

this last equality holding because  $\dim_H \mu_n \rightarrow \dim_A E$  and using [2, Theorem 2.7]. Since for all  $\theta$

$$\min\{\dim_A E, 1\} \geq \min\{\dim_H E, 1\} \geq \dim_H \pi_\theta(E),$$

(3.2) follows.  $\square$

It would be nice to replace the condition in Corollary 3.3 on the existence of the sequence of Bernoulli measures by the requirement that  $\dim_A E = \dim_H E$ . However the relationship between the statements ‘ $\dim_A E = \dim_H E$ ’, ‘ $E$  supports an ergodic measure  $\mu$  with  $\dim_H \mu = \dim_A E$ ’ and ‘there exists a sequence  $(\mu_n)$  of Bernoulli measures supported on  $E$  with  $\dim \mu_n \rightarrow \dim_A E$ ’ is not clear. All three statements hold for almost all sets of translation vectors  $(d_1 \cdots d_k)$  whenever each  $A_i$  has  $\|A_i\| < \frac{1}{2}$ , but the relationship between the statements is difficult to understand.

Note that our results require the matrices  $A_i$  to have positive entries, so that the matrices all map the first quadrant into its interior. This assumption is used in various ways and shortens some of the proofs, but it is crucial in Bárány’s proof [2, Proposition 3.3] that projections of self-affine measures are exact dimensional, which in turn is used in the proof of Theorem 4.6 on averages of  $r$ -scale entropies.

**3.1. Examples.** There are many families of measures on self-affine sets for which  $\dim_H \mu = \dim_L \mu$  and hence for which our main theorem holds. In particular, this includes classes of examples presented in Hueter and Lalley [12], Bárány [2] and Falconer and Kempton [9]. The corollary below shows that this situation also arises for almost all sets of translations in the affine maps in the IFS that defines  $E$ .

**Corollary 3.4.** *Let  $\{T_i\}_{i=1}^k$  be an affine IFS (2.1) where the matrices  $A_i$  are strictly positive with  $\|A_i\| < \frac{1}{2}$  for all  $i$ . For each translation vector  $\mathbf{d} = (d_1, \dots, d_k) \in \mathbf{R}^{2k}$  let  $E_{\mathbf{d}}$  be the self-affine attractor thus defined. Then for almost all  $\mathbf{d} \in \mathbf{R}^{2k}$  (in the sense of  $2k$ -dimensional Lebesgue measure)*

$$\dim_H \pi_\theta(E_{\mathbf{d}}) = \min\{\dim_H E_{\mathbf{d}}, 1\} = \min\{\dim_A E_{\mathbf{d}}, 1\}$$

for all  $\theta \in B$  simultaneously (where as before  $B$  is either  $\mathbf{PR}^1$  or is  $\mathbf{PR}^1$  with one angle omitted).

*Proof.* The right-hand equality follows since  $\dim_H E_{\mathbf{d}} = \min\{\dim_A E_{\mathbf{d}}, 2\}$  for almost all  $\mathbf{d}$ , see [4, 22]. For the left-hand equality we will show that for all  $\epsilon > 0$  we can find a Bernoulli measure  $\mu$  on  $\Sigma$  that induces measures  $\mu_{\mathbf{d}}$  on the  $E_{\mathbf{d}}$  with  $\dim_H \mu_{\mathbf{d}} > \dim_H E_{\mathbf{d}} - \epsilon$  for almost all  $\mathbf{d}$ , so that the conclusion will follow from Corollary 3.3.

As the  $A_i$  have strictly positive entries, there is a cone  $K$  strictly inside the upper-right quadrant that is mapped into itself by all finite compositions of the  $A_i$ . This implies, see [21, Lemma 3.1], that there is a number  $c \geq 1$  that depends only on  $K$  such that for all finite words  $\underline{a} = a_1 \dots a_j, \underline{b} = b_1 \dots b_{j'} \in \Sigma^*$ ,

$$\|A_{\underline{a}}A_{\underline{b}}\| \leq \|A_{\underline{a}}\|\|A_{\underline{b}}\| \leq c\|A_{\underline{a}}A_{\underline{b}}\|.$$

Since  $\phi^s(A) = \|A\|^s$  ( $0 \leq s \leq 1$ ) and  $\phi^s(A) = \|A\|^{2-s}(\det A)^{s-1}$  ( $1 \leq s \leq 2$ ), it follows that

$$(3.3) \quad \phi^s(A_{\underline{a}}A_{\underline{b}}) \leq \phi^s(A_{\underline{a}})\phi^s(A_{\underline{b}}) \leq c\phi^s(A_{\underline{a}}A_{\underline{b}}).$$

for all  $s \geq 0$ .

Write  $d_A := \dim_A(A_1, \dots, A_k) \equiv \dim_A E_{\mathbf{d}}$  for the affinity dimension. We may assume that  $0 < d_A \leq 2$  (if  $d_A > 2$  then  $E_{\mathbf{d}}$  has positive plane Lebesgue measure for almost all  $\mathbf{d}$  and the result is clear). Let  $0 < t < d_A$  ( $t$  not an integer). Using a simple estimate of the rate of decrease of  $\phi^s(A_{\underline{a}})$  with  $s$ , see [4], we may choose an  $N \in \mathbf{N}$  sufficiently large to ensure that for some  $0 < \lambda < 1$

$$(3.4) \quad c \frac{\phi^{d_A}(A_{\underline{a}})}{\phi^t(A_{\underline{a}})} < \lambda \quad \text{for all } \underline{a} = a_1 \dots a_N \in \Lambda^N.$$

From the definition of  $d_A$  (2.2) and submultiplicative properties,

$$\inf_{n \in \mathbf{N}} \left( \sum_{|\underline{a}|=n} \phi^{d_A}(A_{\underline{a}}) \right)^{1/n} = 1,$$

so we may choose  $s \geq d_A$  such that

$$(3.5) \quad \sum_{\underline{a} \in \Lambda^N} \phi^s(A_{\underline{a}}) = \left( \sum_{\underline{a} \in \Lambda^N} \phi^s(A_{\underline{a}}) \right)^{1/N} = 1$$

using the continuity of these sums in  $s$ .

Let  $\Sigma_N^* = \bigcup_{j=0}^\infty \Lambda^{jN}$ . We define a Bernoulli measure  $\mu$  on  $\Sigma$  by specifying  $\mu$  on cylinders defined by words in  $\Sigma_N^*$ :

$$\mu([a_1 \dots a_n]) = \phi^s(A_{a_1})\phi^s(A_{a_2}) \dots \phi^s(A_{a_n}) \quad (n \in \mathbf{N}, a_i \in \Lambda^N).$$

By (3.5)  $\mu$  defines a Borel probability measure on  $\Sigma = (\Lambda^N)^{\mathbf{N}}$ . Using (3.3) and (3.4)

$$(3.6) \quad \frac{\mu([\underline{a}_1 \cdots \underline{a}_n])}{\phi^t(A_{\underline{a}_1 \cdots \underline{a}_n})} \leq \frac{\phi^s(A_{\underline{a}_1}) \cdots \phi^s(A_{\underline{a}_n})}{c^{1-n} \phi^t(A_{\underline{a}_1}) \cdots \phi^t(A_{\underline{a}_n})} \leq \frac{c^{n-1} \phi^{d_A}(A_{\underline{a}_1}) \cdots \phi^{d_A}(A_{\underline{a}_n})}{\phi^t(A_{\underline{a}_1}) \cdots \phi^t(A_{\underline{a}_n})} \leq c^{-1} \lambda^n.$$

We now proceed as in the proof of [4, Theorem 5.3]. We represent points of  $E_{\mathbf{d}}$  in the usual way by  $x_{\mathbf{d}}(\underline{a}) = \bigcap_{i=0}^{\infty} T_{\underline{a}|i}(D)$  for  $\underline{a} \in \Sigma$ . We write  $\mu_{\mathbf{d}}$  for the push-down of  $\mu$  onto  $E_{\mathbf{d}}$ , given by  $\mu_{\mathbf{d}}(F) = \mu\{\underline{a} : x_{\mathbf{d}}(\underline{a}) \in F\}$ . We write  $\underline{a} \wedge \underline{b}$  for the common initial word of  $\underline{a}, \underline{b} \in \Sigma$ . As in [4, Theorem 5.3] we may bound the energy integral of  $\mu_{\mathbf{d}}$  integrated over some disc  $B \subset \mathbf{R}^{2k}$  by

$$\begin{aligned} & \int_{\mathbf{d} \in B} \int_{x \in E_{\mathbf{d}}} \int_{y \in E_{\mathbf{d}}} \frac{d\mu_{\mathbf{d}}(x) d\mu_{\mathbf{d}}(y) d\mathbf{d}}{|x - y|^t} = \int_{\mathbf{d} \in B} \int_{\underline{a} \in \Sigma} \int_{\underline{b} \in \Sigma} \frac{d\mu(\underline{a}) d\mu(\underline{b}) d\mathbf{d}}{|x_{\mathbf{d}}(\underline{a}) - x_{\mathbf{d}}(\underline{b})|^t} \\ & \leq c_1 \int_{\underline{a} \in \Sigma} \int_{\underline{b} \in \Sigma} \phi^t(A_{\underline{a} \wedge \underline{b}})^{-1} d\mu(\underline{a}) d\mu(\underline{b}) \leq c_1 \sum_{\underline{p} \in \Sigma^*} \phi^t(A_{\underline{p}})^{-1} \mu([\underline{p}])^2 \\ & \leq c_1 c_2 c_3 \sum_{\underline{q} \in \Sigma_N^*} \phi^t(A_{\underline{q}})^{-1} \mu([\underline{q}])^2. \end{aligned}$$

For the last inequality we have regarded each word  $\underline{p} \in \Sigma^*$  as  $\underline{p} = \underline{q}q_1 \cdots q_j$  with  $\underline{q} \in \Sigma_N^*$  and  $0 \leq j \leq N - 1$ , and used that  $\phi^t(A_{\underline{p}})^{-1} \mu([\underline{p}])^2 \leq c_2 \phi^t(A_{\underline{q}})^{-1} \mu([\underline{q}])^2$  for some  $c_2$  independent of  $\underline{p}$ , since the summands increase by a bounded factor on adding each single letter to a word  $\underline{q}$ . The constant  $c_3 = \sum_{i=0}^{N-1} k^i$  is the number of words  $\underline{p} \in \Sigma^*$  for which the summands are estimated by the summand of each  $\underline{q} \in \Sigma_N^*$ . Writing  $c_4 = c_1 c_2 c_3$  and using (3.6),

$$\begin{aligned} & \int_{\mathbf{d} \in B} \int_{x \in E_{\mathbf{d}}} \int_{y \in E_{\mathbf{d}}} \frac{d\mu_{\mathbf{d}}(x) d\mu_{\mathbf{d}}(y) d\mathbf{d}}{|x - y|^t} \leq c_4 \sum_{j=0}^{\infty} \sum_{\underline{q} \in \Lambda^{jN}} \phi^t(A_{\underline{q}})^{-1} \mu([\underline{q}]) \mu([\underline{q}]) \\ & \leq c_4 c^{-1} \sum_{j=0}^{\infty} \lambda^j \sum_{\underline{q} \in \Lambda^{jN}} \mu([\underline{q}]) = c_4 c^{-1} \sum_{j=0}^{\infty} \lambda^j < \infty. \end{aligned}$$

We conclude that for almost all  $\mathbf{d}$ ,  $\int_{x \in E_{\mathbf{d}}} \int_{y \in E_{\mathbf{d}}} |x - y|^{-t} d\mu_{\mathbf{d}}(x) d\mu_{\mathbf{d}}(y) < \infty$  implying that for  $\mu_{\mathbf{d}}$ -almost all  $x \in E_{\mathbf{d}}$  there is a constant  $c_0$  such that  $\mu_{\mathbf{d}}(B(x, r)) \leq c_0 r^t$  for all  $r > 0$ , so  $\dim \mu_{\mathbf{d}} \geq t$ . This is the case for all  $t < d_A$ , as required.  $\square$

### 4. Proof of the main theorem

In this section we prove Theorem 3.1. There are two intermediate results which have relatively technical proofs whose main ideas are already well-understood in other work. Since these proofs are quite long and somewhat tangential to the thrust of our argument, we defer them to the appendix.

We first recall a proposition from [9] on the dynamics of projections of self-affine sets. Let  $\pi_{\theta} : D \rightarrow [-1, 1]$  denote orthogonal projection in direction  $\theta$  onto the diameter of  $D$  at angle perpendicular to  $\theta$ , where we identify this diameter (of length 2) with the line  $[-1, 1]$ .

**Proposition 4.1.** *For each  $i \in \{1, \dots, k\}, \theta \in \mathbf{PR}^1$  there is a well defined linear contraction  $f_{i,\theta} : [-1, 1] \rightarrow [-1, 1]$  given by*

$$f_{i,\theta} = \pi_{\theta} \circ T_i \circ \pi_{\phi_i(\theta)}^{-1}$$

such that  $\pi_{\theta}(T_i(A)) = f_{i,\theta}(\pi_{\phi_i(\theta)}(A))$  for all Borel sets  $A \subset D$ .

This identity represents  $\pi_\theta(E)$  as a union of scaled-down copies of projections of  $E$  in other directions.

We let  $\Sigma^*$  denote the space of all finite words made by concatenating the letters  $\{1, \dots, k\}$ . We think of  $\Sigma^*$  as a tree, letting words  $a_1 \dots a_{n+1}$  be children of the parent node  $a_1 \dots a_n$ . Given a direction  $\theta$  in which we want to project  $E$ , we define a length function  $|\cdot|_\theta$  on the tree  $\Sigma^*$  by declaring the length of the path from the root to the node  $a_1 \dots a_n$  to be

$$|a_1 \dots a_n|_\theta = -\log_2 \left( \frac{|\pi_\theta(D_{a_1 \dots a_n})|}{2} \right).$$

Since  $|\pi_\theta(D)| = 2$ , the division by 2 inside the log is necessary to ensure that the length of the empty word is zero.

**Lemma 4.2.** *The length function  $|\cdot|_\theta$  satisfies*

$$|a_1 \dots a_{n+1}|_\theta - |a_1 \dots a_n|_\theta = |a_{n+1}|_{\phi_{a_n \dots a_1}(\theta)}.$$

*Proof.* Identically,

$$\log_2 \left( \frac{|\pi_\theta(D_{a_1 \dots a_{n+1}})|}{2} \right) = \log_2 \left( \frac{|\pi_\theta(D_{a_1 \dots a_n})|}{2} \right) + \log_2 \left( \frac{|\pi_\theta(D_{a_1 \dots a_{n+1}})|}{|\pi_\theta(D_{a_1 \dots a_n})|} \right).$$

Applying the linear map  $T_{a_1 \dots a_n}^{-1}$  to  $D_{a_1 \dots a_n}$  and using Proposition 4.1,

$$\log_2 \left( \frac{|\pi_\theta(D_{a_1 \dots a_{n+1}})|}{|\pi_\theta(D_{a_1 \dots a_n})|} \right) = \log_2 \left( \frac{|\pi_{\phi_{a_n \dots a_1}(\theta)}(D_{a_{n+1}})|/2}{|\pi_{\phi_{a_n \dots a_1}(\theta)}(D)|/2} \right) = \log_2 \left( \frac{|\pi_{\phi_{a_n \dots a_1}(\theta)}(D_{a_{n+1}})|}{2} \right)$$

since  $|\pi_\theta(D)|$ , the length of the projection of the unit disk, is equal to 2 for all  $\theta$ .  $\square$

Let  $N \in \mathbf{N}$  be fixed. For each  $\underline{a} \in \Sigma$ ,  $\theta \in \mathbf{PR}^1$  and  $j \in \mathbf{N}$  we define  $n_j = n_j(\underline{a}, N, \theta)$  to be the natural number satisfying

$$(4.1) \quad |a_1 \dots a_{n_j-1}|_\theta < Nj \leq |a_1 \dots a_{n_j}|_\theta.$$

We define the infinite tree  $\Sigma_N = \Sigma_N(\theta) \subset (\Sigma^*)^\mathbf{N}$  to be the tree with nodes at level  $j$  labeled by words  $a_1 \dots a_{n_j}$ , and edges between level  $j - 1$  nodes and level  $j$  nodes labelled by  $a_{n_{j-1}+1} \dots a_{n_j}$ .

A word  $\underline{a} \in \Sigma$  can also be regarded as an element of  $\Sigma_N$  in which the first symbol is  $a_1 \dots a_{n_1}$ , the second is  $a_{n_1+1} \dots a_{n_2}$  etc. When the distinction between  $\Sigma$  and  $\Sigma_N$  is important we specify which space  $\underline{a}$  is in. The measure  $\mu$  on  $\Sigma$  may also be regarded as a measure on  $\Sigma_N$ , which we also denote by  $\mu$ .

We consider the metric  $d_N$  on the tree  $\Sigma_N$  given by

$$d_N(\underline{a}, \underline{b}) = 2^{-\max\{j: a_1 \dots a_{n_j} = b_1 \dots b_{n_j}\}}.$$

This metric depends on  $\theta$  through the definition of the sequences  $a_{n_j}, b_{n_j}$ . For all sequences  $\underline{a}, \underline{b} \in \Sigma$ , and the corresponding sequences in  $\Sigma_N$ ,

$$|\pi_\theta(\underline{a}) - \pi_\theta(\underline{b})| \leq |\pi_\theta(D_{\underline{a} \wedge \underline{b}})| \leq d_N(\underline{a}, \underline{b})$$

for all  $N$ . In particular the map  $\pi_\theta: (\Sigma_N, d_N) \mapsto [-1, 1]$  is Lipschitz.

Let  $\mu_{[a_1 \dots a_n]} := \mu_{[a_1 \dots a_n]} / \mu[a_1 \dots a_n]$ . We state a variant of results of Hochman and Shmerkin [13] that we will apply to the projections.

**Theorem 4.3.** *If*

$$(4.2) \quad \lim_{N \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{N} \frac{1}{\log 2} \frac{1}{n} \sum_{i=1}^n H_{2^{-(i+1)N}}(\pi_\theta(\mu_{[a_1 \dots a_{n_i}]})) \geq C$$

for  $\mu$ -almost every  $\underline{a} \in \Sigma_N$ , then  $\dim_H \pi_\theta(\mu) \geq C$ .



This statement is essentially contained in Theorems 4.4 and 5.4 of [13]. However, a little work is required to translate those theorems to our setting and the proof is given in the appendix.

The rest of the proof of Theorem 3.1 depends on studying the limits in (4.2). Let

$$\lambda_{\max} := \max\{|a_i|_{\theta} : i \in \{1, \dots, k\}, \theta \in \mathbf{PR}^1\}.$$

**Lemma 4.4.** For all  $\underline{a} \in \Sigma$ , all  $i, N \in \mathbf{N}$  and all  $\theta \in \mathbf{PR}^1$ ,

$$(4.3) \quad H_{2^{-(i+1)N}}(\pi_{\theta}(\mu_{[a_1 \dots a_{n_i}]})) \geq H_{2^{\lambda_{\max}} 2^{-N}}(\pi_{\phi_{a_{n_i} \dots a_1}(\theta)}(\mu)).$$

*Proof.* Note that

$$\pi_{\theta}(\mu_{[a_1 \dots a_{n_i}]}) = S(\pi_{\phi_{a_{n_i} \dots a_1}(\theta)}(\mu)),$$

where  $S$  is the map on measures on  $[-1, 1]$  induced by a linear map on  $[-1, 1]$  which contracts by a factor  $2^{-|a_1 \dots a_{n_i}|_{\theta}}$ . An entropy  $H_{\lambda}(\tau)$  is unchanged by rescaling the measure  $\tau$  provided that we rescale the parameter  $\lambda$  by the same amount. From the definitions of  $\lambda_{\max}$  and  $n_k$ ,

$$|a_1 \dots a_{n_i}|_{\theta} \in [iN, iN + \lambda_{\max}].$$

Then

$$H_{2^{-(i+1)N}}(\pi_{\theta}(\mu_{[a_1 \dots a_{n_i}]})) = H_{\frac{2^{-(i+1)N}}{2^{-|a_1 \dots a_{n_i}|_{\theta}}}}(\pi_{\phi_{a_{n_i} \dots a_1}(\theta)}(\mu)) = H_{\rho \cdot 2^{-N}}(\pi_{\phi_{a_{n_i} \dots a_1}(\theta)}(\mu))$$

where the ‘error’  $\rho$  is given by

$$\rho = \frac{2^{iN}}{2^{-|a_1 \dots a_{n_i}|_{\theta}}} \in [1, 2^{\lambda_{\max}}].$$

Since the entropy  $H_{\lambda}(\mu)$  is monotone decreasing in  $\lambda$ , we obtain (4.3). □

We need to pass from equidistribution results with respect to the sequence  $(\phi_{a_n \dots a_1}(\theta))_{n=1}^{\infty}$  to corresponding results for the subsequence  $(\phi_{a_{n_i} \dots a_1}(\theta))_{i=1}^{\infty}$ .

**Proposition 4.5.** There exists a measure  $\nu_F$  on  $\mathbf{PR}^1$  which is equivalent to  $\mu_F$  and such that for all  $\theta \in B$ ,  $N \in \mathbf{N}$  and  $\mu$ -almost all  $\underline{a} \in \Sigma$ , the sequence  $(\phi_{a_{n_i} \dots a_1}(\theta))_{i=1}^{\infty}$  equidistributes with respect to  $\nu_F$ , i.e., for all intervals  $A \subset \mathbf{PR}^1$  we have

$$\lim_{I \rightarrow \infty} \frac{1}{I} \left| \left\{ i \in \{1, \dots, I\} : \phi_{a_{n_i} \dots a_1}(\theta) \in A \right\} \right| = \nu_F(A).$$

The proof of this proposition is relatively long but involves only standard ergodic theory so is postponed to the appendix.

We now study the sums in (4.2).

**Theorem 4.6.** For all  $\theta \in B$  and  $\mu$ -almost every  $\underline{a} \in \Sigma$ ,

$$(4.4) \quad \lim_{N \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{N \log 2} \frac{1}{n} \sum_{i=1}^n H_{2^{-(i+1)N}}(\pi_{\theta}(\mu_{[a_1 \dots a_{n_i}]})) \geq \beta(\mu).$$

*Proof.* By Lemma 4.4, for  $\theta \in B$  and  $\mu$  almost every  $\underline{a} \in \Sigma$ ,

$$(4.5) \quad \begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{N \log 2} \frac{1}{n} \sum_{i=1}^n H_{2^{-(i+1)N}}(\pi_{\theta}(\mu_{[a_1 \dots a_{n_i}]})) \\ & \geq \liminf_{n \rightarrow \infty} \frac{1}{N \log 2} \frac{1}{n} \sum_{i=1}^n H_{2^{-(N-\lambda_{\max})}}(\pi_{\phi_{a_{n_i} \dots a_1}(\theta)}(\mu)) \\ & \geq \frac{1}{N \log 2} \int_{\mathbf{PR}^1} H_{2^{-(N-\lambda_{\max})}}(\pi_{\alpha}(\mu)) d\nu_F(\alpha). \end{aligned}$$

where the second inequality holds because the function

$$\alpha \mapsto H_{2^{-N}}(\pi_\alpha(\mu))$$

is lower semi-continuous in  $\alpha$  and the sequence  $(\phi_{a_{n_i} \dots a_1}(\theta))_{i=1}^\infty$  equidistributes with respect to  $\nu_F$  by Proposition 4.5.

For  $\mu_F$  almost every  $\alpha$  the measure  $\pi_\alpha(\mu)$  is exact dimensional with dimension  $\beta(\mu)$ , and since  $\mu_F$  and  $\nu_F$  are equivalent the same is true for  $\nu_F$  almost all  $\alpha$ . By (2.7)

$$(4.6) \quad \lim_{N \rightarrow \infty} \frac{1}{(N - \lambda_{\max}) \log 2} H_{2^{-(N - \lambda_{\max})}}(\pi_\alpha(\mu)) = \beta(\mu)$$

for  $\nu_F$  almost every  $\alpha$ .

Noting that  $\lim_{N \rightarrow \infty} \frac{N}{N - \lambda_{\max}} = 1$ , we may integrate (4.6) over  $\alpha$  to get

$$\int_{\mathbf{PR}^1} \lim_{N \rightarrow \infty} \frac{1}{N \log 2} H_{2^{-(N - \lambda_{\max})}}(\pi_\alpha(\mu)) d\nu_F(\alpha) = \beta(\mu).$$

Since this integrand is bounded, we may interchange the limit and integral by the dominated convergence theorem. Combining this with inequality (4.5) gives (4.4).  $\square$

Putting together Theorems 4.3 and 4.6 we conclude that, for all  $\theta \in B$ ,

$$\dim_H \pi_\theta(\mu) \geq \lim_{N \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{N \log 2} \frac{1}{n} \sum_{i=1}^n H_{2^{-(i+1)N}}(\pi_\theta(\mu_{[a_1 \dots a_{n_i}]})) \geq \beta(\mu)$$

which completes the proof of Theorem 3.1.

### 5. Appendix: Technical proofs

**5.1. Proof of Theorem 4.3.** We use the terms ‘tree morphism’ and ‘faithful map’ as defined in [13]: essentially a tree morphism is a map from one sequence space to another which preserves the metric and the structure of cylinder sets, and a faithful map is a map (in our case from a sequence space to  $[-1, 1]$ ) which does not distort dimension too much. The interested reader can find technical definitions of these terms, along with many lemmas of Hochman and Shmerkin which we use but do not write out in full, in [13].

*Proof.* First we define a space  $Y_N$  and split the projection  $\pi_\theta: \Sigma_N \rightarrow [-1, 1]$  into a tree-morphism  $g_N: \Sigma_N \rightarrow Y_N$  which is easier to deal with than  $\pi_\theta$  itself, and a faithful map  $h_N: Y_N \rightarrow [-1, 1]$ .

Let  $Y_N := \{-2^{N+1}, \dots, 2^{N+1} - 1\}^{\mathbf{N}}$  be equipped with metric  $d_N$ . Associate to each  $i \in \{-2^{N+1}, \dots, 2^{N+1} - 1\}$  an interval  $I_i$  of length  $2^{-N}$  with left endpoint at  $\frac{i}{2^{N+1}}$ . This gives an overlapping covering of  $[-1, 1]$ , with each point  $x \in [0, 1]$  contained in either two or three intervals  $I_i$ , see Figure 1.

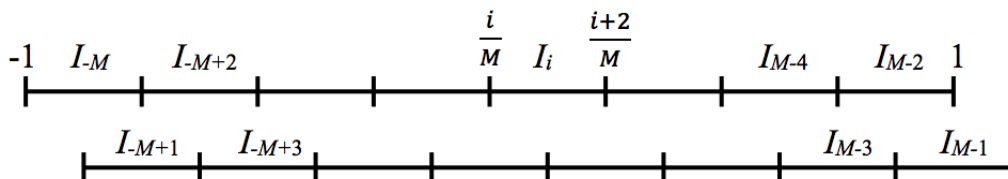


Figure 1. The intervals  $I_i$ , where  $M = 2^{N+1}$ .

Let  $S_i: [-1, 1] \rightarrow [-1, 1]$  be the linear contraction mapping  $[-1, 1]$  onto  $I_i$ , and let  $h_N: Y_N \rightarrow [-1, 1]$  be given by

$$h_N(\underline{a}) = \lim_{n \rightarrow \infty} S_{a_1} \circ \dots \circ S_{a_n}(0) \quad (\underline{a} \in \Sigma).$$

The map  $h_N$  is 3-faithful in the sense of [13] since for each point  $x$  in each interval  $h_N[a_1 \dots a_n]$  there are at most three values of  $a_{n+1} \in \{0, \dots, 2^{N+1} - 1\}$  for which  $x \in h_N[a_1 \dots a_{n+1}]$ . By [13, Proposition 5.2] there exists a constant  $C_2$ , independent of  $N$ , such that for every measure  $\nu$  on  $Y_N$

$$(5.1) \quad \dim_H \nu - \frac{C_2}{N \log 2} \leq \dim_H h_N(\nu).$$

With the standard binary coding of  $[-1, 1]$  it is not the case that for any interval  $A \subset [-1, 1]$  with  $|A| < 2^{-(j+1)}$  one has that  $A$  is contained in some binary interval of depth  $j$ . However, with the overlapping coding given by the intervals  $I_i$  and the map  $h_N$ , it holds that for any  $A$  with  $|A| < 2^{-N(j+1)}$  there is a choice of  $a_1 \dots a_j$  such that  $A \subset h_N[a_1 \dots a_j]$ . Furthermore, these codings can be iterated in the sense that if  $B \subset A$  has  $|B| < 2^{-N(j+2)}$  then  $B$  is contained in the interval  $h_N[a_1 \dots a_j b_{j+1}]$  for some choice of  $b_{j+1}$ .

We now define a map  $g_N: \Sigma_N \rightarrow Y_N$  which maps depth- $n$  cylinders in  $\Sigma_N$  to depth- $n$  cylinders in  $Y_N$ . The map  $g_N$  is defined iteratively. Given a depth-2 cylinder  $[a_1 \dots a_{n_2}] \subset \Sigma_N$ , there exists at least one depth-1 cylinder  $[b_1] \subset Y_N$  such that  $\pi_\theta(D_{a_1 \dots a_{n_2}}) \subset h_N([b_1])$ , since  $|\pi_\theta(D_{a_1 \dots a_{n_2}})| < 2^{-N}$ . We choose such a cylinder  $[b_1]$ .

Now assume we have already chosen a depth- $(j - 1)$  cylinder  $[b_1 \dots b_{j-1}]$  corresponding to word  $a_1 \dots a_{n_j}$ . Then we may choose for each depth- $(j + 1)$  cylinder  $[a_1 \dots a_{n_{j+1}}] \subset [a_1 \dots a_{n_j}]$  a letter  $b_j$  such that

$$\pi_\theta(D_{a_1 \dots a_{n_{j+1}}}) \subset h_N([b_1 \dots b_j]).$$

This process defines a tree morphism  $g_N: \Sigma_N \rightarrow Y_N$  such that  $\pi_\theta: \Sigma_N \rightarrow [-1, 1]$  can be written  $\pi_\theta = h_N \circ g_N$ . Since  $h_N$  is faithful and  $\pi_\theta = h_N \circ g_N$ ,

$$\left| H_{2^{-(i+1)N}}(\pi_\theta(\mu_{[a_1 \dots a_{n_i}]})) - H_{2^{-(i+1)N}}(g_N(\mu_{[a_1 \dots a_{n_i}]})) \right|$$

is bounded above by some constant which is independent of  $N$ . In particular, inequality (4.2) holds if and only if

$$\lim_{N \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{N \log 2} \frac{1}{n} \sum_{i=1}^n H_{2^{-(i+1)N}}(g_N(\mu_{[a_1 \dots a_{n_i}]})) \geq C.$$

But since  $g_N$  is a tree morphism, [13, Theorem 4.4] gives

$$\lim_{N \rightarrow \infty} \dim_H g_N(\mu) \geq C,$$

so setting  $\nu = g_N(\mu)$  in (5.1),

$$\dim_H \pi_\theta(\mu) \geq C. \quad \square$$

**5.2. Proof of Proposition 4.5.** Proposition 4.5 involves transferring an equidistribution result for sequences  $(\phi_{a_n \dots a_1}(\theta))_{n=1}^\infty$  to a different equidistribution result for the subsequences  $(\phi_{a_{n_i} \dots a_1}(\theta))_{i=1}^\infty$ . To do this we first need to encode the subsequences dynamically, which can be done by building a suspension flow based around the maps  $\phi_i$ . In order to pick out the correct terms  $n_i$  we introduce a notion of time for our flow based on the numbers  $|a_1|_\theta$ . This is a standard approach in ergodic theory when one is interested in looking at a system with a different notion of time than is given by just counting iterations of a map.

Secondly, we need to ask whether the time- $N$  map associated to this suspension flow is ergodic. There are classical results in ergodic theory which deal with this question, see for example [20]. We verify that our time- $N$  map is indeed ergodic, allowing us to conclude our equidistribution result. The fact that we need Proposition 4.5 to hold for all  $\theta \in \mathcal{Q}_2$ , and not just almost all  $\theta$ , also requires us to prove some extra continuity results on the roof function of our flow.

We begin by defining a map  $T: \Sigma \times \mathbf{PR}^1 \rightarrow \Sigma \times \mathbf{PR}^1$  by

$$T(\underline{a}, \theta) = (\sigma(\underline{a}), \phi_{a_1}(\theta)).$$

We can map the two-sided full shift  $\Sigma^\pm$  onto  $\Sigma \times \mathbf{PR}^1$  by mapping two-sided sequence  $\underline{a}^\pm$  to  $(a_1 a_2 a_3 \cdots, \lim_{n \rightarrow \infty} \phi_{a_0 \dots a_n}(0))$ . Then  $T$  is a factor of the right shift  $\sigma^{-1}$  under this factor map, and the two-sided extension  $\bar{\mu}$  of  $\mu$  is mapped to  $\mu \times \mu_F$ . Since factors of mixing dynamical systems are mixing, the system  $(\Sigma \times \mathbf{PR}^1, \mu \times \mu_F, T)$  is mixing, see [23] for a proof of this and other basic statements in ergodic theory.

The map  $T$  helps us understand the sequence  $(\phi_{a_{n_i} \dots a_1}(\theta))_{i=1}^\infty$ , but we need to know about the sequence  $n_i = n_i(\underline{a}, \theta, N)$ . For this we build a suspension flow over  $T$  by introducing a notion of time. Consider the space

$$W := \{(\underline{a}, \theta, t) \in \Sigma \times \mathbf{PR}^1 \times \mathbf{R}: 0 \leq t \leq |a_1|_\theta\}$$

where the points  $(\underline{a}, \theta, |a_1|_\theta)$  and  $(\sigma(\underline{a}), \phi_{a_1}(\theta), 0)$  are identified. Define a suspension semiflow  $\psi$  on  $W$  by letting

$$\psi_s(\underline{a}, \theta, t) = (\underline{a}, \theta, t + s)$$

for  $0 \leq t + s \leq |a_1|_\theta$ , and extending this to a well-defined semiflow for all  $s > 0$ , using the identification  $(\underline{a}, \theta, |a_1|_\theta) = (\sigma(\underline{a}), \phi_{a_1}(\theta), 0)$ . This flow preserves the measure  $(\mu \times \mu_F \times \mathcal{L})|_W$ , where  $\mathcal{L}$  is Lebesgue measure. Since suspension semiflows over ergodic maps are ergodic,  $\psi_s$  is ergodic.

We now prove some uniform continuity lemmas involving  $\theta, \theta' \in \mathcal{Q}_2$ , which also give non-uniform continuity for  $\theta, \theta' \in B$ , since all  $\theta \in B$  are eventually mapped into  $\mathcal{Q}_2$  by compositions of the  $\phi_i$ . We first show that small variations of  $\theta$  have little effect on the time taken to flow through the base a given number of times.

**Lemma 5.1.** *There exists a constant  $C$  such that for all  $\theta, \theta' \in \mathcal{Q}_2$ ,  $n \in \mathbf{N}$  and  $\underline{a} \in \Sigma$ ,*

$$||a_1 \cdots a_n|_\theta - |a_1 \cdots a_n|_{\theta'}| < C|\theta - \theta'|.$$

*Proof.* For each  $i \in \{1, \dots, k\}$ , the map  $\theta \mapsto |a_i|_\theta$  is differentiable with derivative bounded above by 2, since  $\sin$  and  $\cos$  have derivatives bounded by 1. Furthermore, there exists a constant  $0 < \rho < 1$  such that the maps  $\phi_i$  restricted to  $\mathcal{Q}_2$  are strict contractions with derivative bounded above by  $\rho$ . Then

$$\begin{aligned} ||a_1 \cdots a_n|_\theta - |a_1 \cdots a_n|_{\theta'}| &\leq \sum_{i=1}^n ||a_i|_{\phi_{a_{n-1} \dots a_1}(\theta)} - |a_i|_{\phi_{a_{n-1} \dots a_1}(\theta')}| \\ &\leq (2 \max_{i, \theta''} |a_i|_{\theta''}) \sum_{i=1}^n |\phi_{a_{n-1} \dots a_1}(\theta) - \phi_{a_{n-1} \dots a_1}(\theta')| \\ &\leq (2 \max_{i, \theta''} |a_i|_{\theta''}) |\theta - \theta'| \left( \sum_{i=1}^n \rho^{i-1} \right) < C|\theta - \theta'| \end{aligned}$$

for all  $\theta, \theta' \in \mathcal{Q}_2$ . □

**Lemma 5.2.** *Suppose that the orbit under  $\psi$  of point  $(\underline{a}, \theta, t)$  equidistributes with respect to the measure  $(\mu \times \mu_F \times \mathcal{L})|_W$ . Then the same is true of  $(\underline{a}, \theta', t')$  for all  $\theta' \in \mathcal{Q}_2$  and  $t' \in [0, |a_1|_{\theta'}]$ .*

*Proof.* The distance  $|T^n(\underline{a}, \theta) - T^n(\underline{a}, \theta')| \rightarrow 0$  as  $n \rightarrow \infty$  since the maps  $\phi_i$  are strict contractions on  $\mathcal{Q}_2$ . All we need to check is that there is not too much time distortion when we replace the transformation  $T$  with the flow  $\psi$ . But this has been dealt with by the previous lemma.  $\square$

Given a flow  $\psi_s$  on  $W$ , the *time- $N$  map* is the map  $\psi_N$  on  $W$  regarded as a discrete time dynamical system.

**Lemma 5.3.** *The time- $N$  map  $\psi_N: W \rightarrow W$  preserves the measure  $(\mu \times \mu_F \times \mathcal{L})|_W$  and  $(W, \psi_N, (\mu \times \mu_F \times \mathcal{L})|_W)$  is ergodic.*

*Proof.* First we deal with the case when there exists an irrational number  $c > 0$  such that every periodic orbit of the suspension flow has period equal to an integer multiple of  $c$ . Then it was argued in [20, Proposition 5], using the Livschitz Theorem, that the roof function of our flow is cohomologous to a function which always takes values equal to an integer multiple of  $c$ . We add this coboundary, which has no effect on the dynamics of the flow or the time  $N$  map, and assume now that our roof function always takes value equal to an integer multiple of  $c$ . Now using the fact that the base map  $T$  of our suspension flow is mixing, coupled with the fact that the map  $x \mapsto x + N \pmod{c}$  is ergodic, it follows that the map  $\psi_N$  is ergodic.

In the case that there is a rational number  $c > 0$  such that every periodic orbit has period equal to an integer multiple of  $c$ , we replace  $\log_2$  with  $\log_3$  in the definition of our distance function  $|\cdot|_{\theta}$  which makes the roof function. Then the corresponding constant for our new suspension flow is irrational, and we apply the previous argument in this setting.

Finally we deal with the case that there exists no  $c > 0$  such that every periodic orbit of the suspension flow has period equal to an integer multiple of  $c$ . In [20, Proposition 5] it is shown that this ensures that the flow is weak mixing. But if the suspension flow  $\psi$  is weak mixing then the corresponding time- $N$  map  $\psi_N$  is ergodic as required.  $\square$

Finally we complete the proof of Proposition 4.5. Let  $\pi_2: W \rightarrow \mathbf{PR}^1$  be given by  $\pi_2(\underline{a}, \theta, t) = \theta$ , and let the measure  $\nu_F$  on  $\mathbf{PR}^1$  be  $\nu_F = ((\mu \times \mu_F \times \mathcal{L})|_W) \circ \pi_2^{-1}$ . This measure is equivalent to  $\mu_F$ . If  $\psi_N(\underline{a}, \theta, 0)$  equidistributes with respect to  $(\mu \times \mu_F \times \mathcal{L})$  then the sequence  $(\phi_{a_{n_i} \dots a_1}(\theta))_{i=1}^{\infty}$  equidistributes with respect to  $\nu_F$ . This is because, from the definition of the sequence  $a_{n_i}$ ,

$$\phi_{a_{n_i-1} \dots a_1}(\theta) = \pi_2(\psi_{iN}(\underline{a}, \theta, 0)).$$

Then it is enough to prove that for all  $\theta \in B$ , for  $\mu$ -almost all  $\underline{a}$ , the sequence  $(\psi_{iN}(\underline{a}, \theta, 0))_{i=0}^{\infty}$  equidistributes with respect to  $\nu_F$ . We have already argued that this holds for  $(\mu \times \mu_F)$ -almost all pairs  $(\underline{a}, \theta)$ , so the extension to all  $\theta$  follows readily from Lemma 5.2 and the fact that for all  $\theta \in B$ , for  $\mu$ -almost every  $\underline{a}$ , the orbit  $(\phi_{a_n \dots a_1}(\theta))$  gets arbitrarily close to some  $\mu_F$ -typical point of  $\mathbf{PR}^1$ .

## References

- [1] ALMARZA, J. I.: CP-chains and dimension preservation for projections of  $(\times m, \times n)$ -invariant Gibbs measures. - *Adv. Math.* 304, 2017, 227–265.
- [2] BÁRÁNY, B.: On the Ledrappier–Young formula for self-affine measures. - *Math. Proc. Cambridge Philos. Soc.* 159:3, 2015, 405–432.
- [3] BÁRÁNY, B., M. POLLICOTT, and K. SIMON: Stationary measures for projective transformations: the Blackwell and Furstenberg measures. - *J. Stat. Phys.* 148:3, 2012, 393–421.
- [4] FALCONER, K. J.: The Hausdorff dimension of self-affine fractals. - *Math. Proc. Cambridge Philos. Soc.* 103:2, 1988, 339–350.
- [5] FALCONER, K.: Dimensions of self-affine sets: A survey. - In: *Further developments in fractals and related fields*, Trends Math., Birkhäuser/Springer, New York, 2013, 115–134.
- [6] FALCONER, K.: *Fractal geometry: Mathematical foundations and applications*. - John Wiley & Sons, Ltd., Chichester, third edition, 2014.
- [7] FALCONER, K., J. FRASER, and X. JIN: Sixty years of fractal projections. - In: *Fractal geometry and stochastics V*, Progr. Probab., Birkhäuser/Springer, New York, 2015, 115–134.
- [8] FALCONER, K. J., and X. JIN: Exact dimensionality and projections of random self-similar measures and sets. - *J. Lond. Math. Soc. (2)* 90:2, 2014, 388–412.
- [9] FALCONER, K., and T. KEMPTON: Planar self-affine sets with equal Hausdorff, box and affinity dimensions. - *Ergodic Theory Dynam. Systems*, 2016 (to appear).
- [10] FARKAS, Á.: Projections of self-similar sets with no separation condition. - *Israel J. Math.* (to appear).
- [11] FERGUSON, A., J. M. FRASER, and T. SAHLSTEN: Scaling scenery of  $(\times m, \times n)$  invariant measures. - *Adv. Math.* 268, 2015, 564–602.
- [12] HUETER, I., and S. P. LALLEY: Falconer’s formula for the Hausdorff dimension of a self-affine set in  $\mathbf{R}^2$ . - *Ergodic Theory Dynam. Systems* 15:1, 1995, 77–97.
- [13] HOCHMAN, M., and P. SHMERKIN: Local entropy averages and projections of fractal measures. - *Ann. of Math. (2)* 175:3, 2012, 1001–1059.
- [14] HUTCHINSON, J. E.: Fractals and self-similarity. - *Indiana Univ. Math. J.* 30:5, 1981, 713–747.
- [15] JORDAN, T., M. POLLICOTT, and K. SIMON: Hausdorff dimension for randomly perturbed self affine attractors. - *Comm. Math. Phys.* 270:2, 2007, 519–544.
- [16] KAUFMAN, R.: On Hausdorff dimension of projections. - *Mathematika* 15, 1968, 153–155.
- [17] MARSTRAND, J. M.: Some fundamental geometrical properties of plane sets of fractional dimensions. - *Proc. London Math. Soc. (3)* 4, 1954, 257–302.
- [18] MATTILA, P.: *Geometry of sets and measures in Euclidean spaces: Fractals and rectifiability*. - Cambridge Stud. Adv. Math. 44, Cambridge Univ. Press, Cambridge, 1995.
- [19] PERES, Y., and P. SHMERKIN: Resonance between Cantor sets. - *Ergodic Theory Dynam. Systems* 29:1, 2009, 201–221.
- [20] QUAS, A., and T. SOO: Weak mixing suspension flows over shifts of finite type are universal. - *J. Mod. Dyn.* 6:4, 2012, 427–449.
- [21] SHMERKIN, P.: Self-affine sets and the continuity of subadditive pressure. - In: *Geometry and analysis of fractals*, Springer Proc. Math. Stat. 88, Springer, Heidelberg, 2014, 325–342.
- [22] SOLOMYAK, B.: Measure and dimension for some fractal families. - *Math. Proc. Cambridge Philos. Soc.* 124:3, 1998, 531–546.
- [23] WALTERS, P.: *An introduction to ergodic theory*. - Grad. Texts in Math. 79, Springer-Verlag, New York-Berlin, 1982.