LIPSCHITZ CONTINUITY OF THE HAUSDORFF DIMENSION OF SELF-AFFINE SPONGES AT SIERPINSKI SPONGES

NUNO LUZIA

Abstract. The Hausdorff dimension of general Sierpinski carpets, [4] and [20], and the generalization on Lalley-Gatzouras carpets, [10], are today well known results, the formulas being obtain via the variational principle for the dimension. We call the multidimensional versions of these carpets Sierpinski sponges and self-affine sponges, respectively. In this paper we show that the Hausdorff dimension of self-affine sponges, defined in $\mathbb{R}^3$, is a Lipschitz continuous function at Sierpinski sponges.

1. Introduction and statements

The dimension theory of $C^{1+\alpha}$ conformal repellers is well understood by means of the thermodynamic formalism introduced by Sinai-Ruelle-Bowen [24], [23], [5] and the famous Bowen’s equation [6], [22]. In particular there is a unique ergodic measure of full dimension which is a Gibbs state.

The dimension theory of non-conformal repellers is still being developed and no such general formalism exists. The computation of Hausdorff dimension of non-conformal fractals began with the fundamental works by Bedford [4] and McMullen [20] on the general Sierpinski carpets, and their generalization [10] on the Lalley-Gatzouras carpets, as they are known today. See also [1] for an extension of the Lalley-Gatzouras carpets. These are self-affine fractals in the plane and there is an ergodic measure of full dimension, in fact Bernoulli (such a measure is, in general, not unique for Lalley-Gatzouras carpets [2]). There are also some non-linear versions of these results, [11], [15], [17] and [19], towards a Dimension Formalism for these kind of non-conformal repellers. In particular, in [15] and [19] we compute the Hausdorff dimension of Non-linear Lalley-Gatzouras carpets which, as the name suggests, are the $C^{1+\alpha}$ non-linear versions of the Lalley-Gatzouras carpets.

Other approaches try to give a formula for the Hausdorff dimension in a generic setting, instead of considering particular families as before. One such example is the famous Falconer’s formula [8] which gives, under some hypotheses, the Hausdorff dimension of the self-affine fractal for ‘almost every’ translation vector-parameters, but does not tell for which parameters it holds (exceptions being the self-affine fractals in the plane considered in [12] and its non-linear versions [10]). In fact these two approaches are quite different, for the self-affine fractals for which Falconer’s formula holds there is coincidence with Hausdorff and box-counting dimensions, as for the Lalley-Gatzouras carpets these two dimensions do not coincide in general [10]. More recently, Hochmann and Rapaport [13] gave a formula for the Hausdorff dimension of self-affine sets in the plane satisfying an irreducibility condition and an exponential separation condition.

The computation of the Hausdorff dimension of non-conformal fractals in $\mathbb{R}^d$, $d > 2$ reveals to be a much more difficult task. There is a natural way of defining the $d$-dimensional versions of general Sierpinski carpets and Lalley-Gatzouras
carpets which we shall call, respectively, Sierpinski sponges and self-affine sponges. Recently, in a major breakthrough, [7] showed that the variational principle for the dimension does not hold, in general, within the class of Baranski sponges (even for $d = 3$), i.e. there is not an ergodic measure of full dimension. They showed that the Hausdorff dimension in that class can be calculated via pseudo-Bernoulli measures. On the other hand, Kenyon and Peres [14] computed the Hausdorff dimension of Sierpinski sponges by proving the variational principle for the dimension and, moreover, there is a unique ergodic measure of full Hausdorff dimension, in fact Bernoulli (see also [9], and see [3] for a random version). We do not know if the variational principle for the dimension holds for self-affine sponges that are close to Sierpinski sponges but we will show that the Hausdorff dimension of self-affine sponges, $d = 3$, is Lipschitz continuous at Sierpinski sponges. Since the variational principle for the dimension holds for Sierpinski sponges, this implies that the variational principle for the dimension almost holds for self-affine sponges close to Sierpinski sponges. We notice that the continuity of the Hausdorff dimension in the class of Baranski sponges was proved in [7], but the class of self-affine sponges considered in this paper need not be Baranski sponges. Also by restricting to the continuity of Hausdorff dimension at Sierpinski sponges we are able to obtain more explicit and quantitative results.

1.1. Sierpinski sponges. We begin by describing the Sierpinski sponges studied in [14] (the multidimensional versions of the general Sierpinski carpets). Let $T_d = \mathbb{R}^d / \mathbb{Z}^d$ be the $d$-dimensional torus and $f : T_d \to T_d$ be given by

$$f(x_1, x_2, ..., x_d) = (l_1 x_1, l_2 x_2, ..., l_d x_d)$$

where $l_1 \geq l_2 \geq ... \geq l_d > 1$ are integers. The grids of hyperplanes

$$\{i/l_1\} \times [0, 1]^{d-1}, i = 0, ..., l_1 - 1$$

$$[0, 1] \times \{i/l_2\} \times [0, 1]^{d-2}, i = 0, ..., l_2 - 1$$

$$\vdots$$

$$[0, 1]^{d-1} \times \{i/l_d\}, i = 0, ..., l_d - 1$$

form a set of boxes each of which is mapped by $f$ onto the entire torus (these boxes are the domains of invertibility of $f$). Now choose some of these boxes and consider the fractal set $\Lambda$ consisting of those points that always remain in these chosen boxes when iterating $f$. Geometrically, $\Lambda$ is the limit (in the Hausdorff metric), or the intersection, of $n$-approximations: the 1-approximation consists of the chosen boxes; the 2-approximation consists in replacing each box of the 1-approximation by a rescaled affine copy of the 1-approximation, resulting in more and smaller boxes; the 3-approximation consists in replacing each box of the 2-approximation by a rescaled affine copy of the 1-approximation, and so on. We say that $\Lambda$ is a Sierpinski sponge, and their Hausdorff dimension was computed in [14] (a formula is given in next section).

1.2. Self-affine sponges. Now we describe the self-affine sponges (the multidimensional versions of Lalley-Gatzouras carpets), which include as a very particular case the Sierpinski sponges. Let $S_1, S_2, ..., S_r$ be contractions of $\mathbb{R}^d$. Then there is a unique nonempty compact set $\Lambda$ of $\mathbb{R}^d$ such that

$$\Lambda = \bigcup_{i=1}^{r} S_i(\Lambda).$$

We will refer to $\Lambda$ as the limit set of the semigroup generated by $S_1, S_2, ..., S_r$. We are going to consider sets $\Lambda$ which are limit sets of the semigroup generated by the
$d$-dimensional mappings $A_{i^1,i^2,...,i^d}$ given by

$$A_{i^1,i^2,...,i^d} = \begin{pmatrix}
a_{i^1,i^2,...,i^d} & 0 & 0 & \cdots & 0 \\
0 & a_{i^1,i^2,...,i^{d-1}} & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & a_{i^1} & 0
\end{pmatrix} x + \begin{pmatrix}
u_{i^1,i^2,...,i^d} \\
u_{i^1,i^2,...,i^{d-1}} \\
\vdots \\
u_{i^1}
\end{pmatrix}$$

for $(i^1, i^2, ..., i^d) \in \mathcal{I}$. Here

$$\mathcal{I} = \{(i^1, i^2, ..., i^d) : 1 \leq i^1 \leq m, 1 \leq i^2 \leq m_{i^1}, 1 \leq i^3 \leq m_{i^1i^2}, ..., 1 \leq i^d \leq m_{i^1i^2...i^{d-1}}\}$$

is a finite index set, and $0 < a_{i^1...i^k} < 1$, $k = 1, ..., d$ satisfy

$$a_{i^1...i^k}^{i^k+1} \leq a_{i^1...i^k}.$$  

Also, for each $(i^1, ..., i^d) \in \mathcal{I}$ and $k \in \{1, ..., d\}$,

$$\sum_{i^k=1}^{m_{i^1...i^{k-1}}} a_{i^1...i^k} \leq 1$$

(with $k = 1$ the end of the sum is $m$) and

$$0 \leq u_{i^1...i^k} < u_{i^1...i^{k+1}} < 1, \quad u_{i^1...i^{k+1}} - u_{i^1...i^k} \geq a_{i^1...i^k},$$

when $k > 1$, $i^k = m_{i^1...i^{k-1}}$ we substitute $u_{i^1...i^{k+1}}$ by 1. These hypotheses guarantee that the boxes

$$R_{i^1...i^d} = A_{i^1...i^d}([0, 1]^d)$$

have interiors that are pairwise disjoint, with edges parallel to the coordinate axes, the box $R_{i^1...i^d}$ having $k^{th}$-edge with length $a_{i^1...i^k}$. Geometrically, $\Lambda$ is constructed like the Sierpinski sponge, with the 1-approximation consisting of the boxes $R_{i^1...i^d}$, the 2-approximation consisting in replacing each box of the 1-approximation by a rescaled affine copy of the 1-approximation, and so on. See Figure 1 for an illustration of the case $d = 3$ (where we used $c_i$, $b_{ij}$, $a_{ijk}$ instead of $a_{i^1}$, $a_{i^1i^2}$, $a_{i^1i^2i^3}$, respectively). When $d = 2$ the sets $\Lambda$ are the Lalley-Gatzouras carpets, and their 1-approximation corresponds to the projection onto the $yz$-plane of Figure 1. In general we say that $\Lambda$ is a self-affine sponge.

Given $p = (p_{i^1...i^{d-1}})$ a collection of non-negative numbers satisfying

$$p_{i^1...i^k} = \sum_{i^{k+1}} p_{i^1...i^k,i^{k+1}}, \quad k = d - 2, d - 3, ..., 1$$

$$\sum_{i^k} p_{i^1} = 1,$$

we define the numbers $\lambda_k(p)$, $k = 1, ..., d - 1$ by

$$\lambda_k(p) = \sum_{i^1...i^k} p_{i^1...i^k} \log p_{i^1...i^k} - \sum_{i^1...i^{k-1}} p_{i^1...i^{k-1}} \log p_{i^1...i^{k-1}},$$

(by convention: $0 \log 0 = 0$; for $k = 1$ the second sum in the numerator is 0) and the number $t(p)$ as the unique real in $[0, 1]$ satisfying

$$\sum_{i^1...i^{d-1}} p_{i^1...i^{d-1}} \log \left(\sum_{i^d} a_{i^1...i^{d-1},i^d}^{t(p)}\right) = 0.$$
The number $\sum_{k=1}^{d-1} \lambda_k(p)$ is the Hausdorff dimension in the hyperplane $x_2...x_d$ of the set of generic points for the distribution $p$; the number $t(p)$ is the Hausdorff dimension of a typical 1-dimensional fibre in the $x_1$-direction relative to the distribution $p$, and is given by a random Moran formula. We will see that the sum of these two numbers is the Hausdorff dimension of a Bernoulli measure supported on $\Lambda$, and so we have the following result.

**Theorem 1.** Let $\Lambda$ be a self-affine sponge. Then

$$\dim_H \Lambda \geq \sup_p \left\{ \sum_{k=1}^{d-1} \lambda_k(p) + t(p) \right\}.$$  \hspace{1cm} (1)

Here $\dim_H \Lambda$ stands for the Hausdorff dimension of a set $\Lambda$.

**Problem.** When does equality hold in (1)?

In other words, when does the variational principle for the dimension hold, in the class of self-affine sponges? When $d = 2$ this corresponds to Lalley-Gatzouras carpets and equality in (1) was proved in [10] (see also [15] for an alternative proof and [18] for a random version). For general $d$ but with $a_{ij}...k = a(k), k = 1,...,d$ this corresponds, essentially, to Sierpinski sponges and equality in (1) was, essentially, proved in [14].

As we know now, by [7], equality in (1) might not hold, in general, even when $d = 3$.

For the sake of clarity we will restrict to $d = 3$ (even though the results in this paper might extend to general $d$). In this case, we use the simpler notation $c_i$ for $a_i$, $i = 1,...,m$ and $b_{ij}$ for $a_{ij}$, $j = 1,...,m_j$, keeping the notation for $a_{ijk}$, $k = 1,...,m_{ij}$ (see Figure 1).

2. $(a,b,c;\varepsilon)$-sponges

In this section $d = 3$.

**Definition 1.** We say that a self-affine sponge is an $(a,b,c;\varepsilon)$-sponge, for some numbers $0 < a \leq b \leq c < 1$, $\varepsilon \geq 0$, if

$$e^{-\varepsilon} \leq \frac{a_{ijk}}{a} \leq e^{\varepsilon}, \quad e^{-\varepsilon} \leq \frac{b_{ij}}{b} \leq e^{\varepsilon}, \quad e^{-\varepsilon} \leq \frac{c_i}{c} \leq e^{\varepsilon},$$
for every \((i, j, k) \in \mathcal{I}\).

The case \(\varepsilon = 0\) corresponds essentially to what we called Sierpinski sponges (even though the numbers \(a, b\) and \(c\) need not be integers), and we will still call Sierpinski sponges to this larger class. As said before, their Hausdorff dimension was, essentially, computed in \([14]\) via the variational principle for the dimension. When we are close to a Sierpinski sponge, say an \((a, b, c; \varepsilon)\)-sponge with \(\varepsilon > 0\) small, we do not know if the variational principle for the dimension holds but we know that the Hausdorff dimension is close to the Hausdorff dimension of the Sierpinski sponge \((a, b, c; 0)\). When we talk about continuity of the Hausdorff dimension of self-affine sponges it is implicit that their alphabet \(\mathcal{I}\) is fixed.

Denote by \(\Lambda_{a,b,c}\) a Sierpinski sponge \((a, b, c; 0)\), and by \(\Lambda_{a,b,c;\varepsilon}\) an \((a, b, c; \varepsilon)\)-sponge.

Let \(\mathcal{J} = \{(i,j) : (i,j,k) \in \mathcal{I} \text{ for some } k\}\). Let
\[
\ell = \min_{(i,j) \in \mathcal{J}} \frac{\log m_{ij}}{-\log a} \quad \text{and} \quad \tau = \max_{(i,j) \in \mathcal{J}} \frac{\log m_{ij}}{-\log a}.
\]

**Theorem 2.** Let \(\Lambda_{a,b,c}\) be a Sierpinski sponge such that \(\ell < 7\). There exists a constant \(C\) (depending only on \(a, b, c\) and \(\mathcal{I}\)) such that for every \(\varepsilon > 0\) sufficiently small
\[
\dim_H \Lambda_{a,b,c;\varepsilon} - C\varepsilon \leq \dim_H \Lambda_{a,b,c;\varepsilon} \leq \dim_H \Lambda_{a,b,c} + C\varepsilon.
\]

**Remark 1.** Going through the proof of Theorem 2 it is possible to give an explicit expression for \(C\), if that is needed for some application, but that is not our purpose in this paper.

The lower estimate in Theorem 2 follows easily from Theorem 1. For the upper estimate in Theorem 2 we will construct a 2-parameter family of Bernoulli measures.

Given \(\Lambda\) a self-affine sponge we write
\[
\text{VP}(\Lambda) = \sup \left\{ \lambda_1(p) + \lambda_2(p) + t(p) \right\}.
\]

Even though we do not know if the variational principle for the dimension holds for \((a, b, c; \varepsilon)\)-sponges, \(\varepsilon > 0\) small, we have the following result.

**Corollary 1.** Let \(\Lambda_{a,b,c;\varepsilon}\) be a Sierpinski sponge such that \(\ell < 7\). There exists a constant \(C\) (depending only on \(a, b, c\) and \(\mathcal{I}\)) such that for every \(\varepsilon > 0\) sufficiently small
\[
\text{VP}(\Lambda_{a,b,c;\varepsilon}) \leq \dim_H \Lambda_{a,b,c;\varepsilon} \leq \text{VP}(\Lambda_{a,b,c;\varepsilon}) + C\varepsilon.
\]

The lower estimate in Corollary 1 follows from Theorem 1, and the upper estimate in Corollary 1 follows from the proof of Theorem 2.

### 3. Proof of Theorem 1

In this part, \(d\) is any integer greater than or equal to 2.

There is a natural symbolic representation associated with our system that we shall describe now. Consider the sequence space \(\Omega = \mathcal{T}^N\). Elements of \(\Omega\) will be represented by \(\omega = (\omega_1, \omega_2, \ldots)\) where \(\omega_n = (i^n_1, \ldots, i^n_d) \in \mathcal{I}\). Given \(\omega \in \Omega\) and \(n \in \mathbb{N}\), let \(\omega(n) = (\omega_1, \omega_2, \ldots, \omega_n)\) and define the cylinder of order \(n\),
\[
C_{\omega(n)} = \{\omega' \in \Omega : \omega'_l = \omega_l, \ l = 1, \ldots, n\}
\]
and the basic box of order \(n\),
\[
R_{\omega(n)} = A_{i^n_1} \circ A_{i^n_2} \circ \cdots \circ A_{i^n_d}([0, 1]^d).
\]
We have that \((R_{\omega(n)})_n\) is a decreasing sequence of closed boxes having \(k^n\)-edge with length \(\prod_{l=1}^d a_{il}^{-1} i_l\). Thus \(\bigcap_{n=1}^\infty R_{\omega(n)}\) consists of a single point which belongs
to $\Lambda$ that we denote by $\chi(\omega)$. This defines a continuous and surjective map $\chi: \Omega \to \Lambda$ which is at most $2^d$ to 1, and only fails to be a homeomorphism when some of the boxes $R_{i_1 \ldots i_d}$ have nonempty intersection.

Let $\lambda(p) = \sum_{k=1} N \Lambda_k(p)$. We shall construct probability measures $\mu_p$ supported on $\Lambda$ with

$$\dim_H \mu_p = \lambda(p) + t(p).$$

This gives what we want because $\dim_H \Lambda \geq \dim_H \mu_p$.

Let $\tilde{\mu}_p$ be the Bernoulli measure on $\Omega$ that assigns to each symbol $(i_1, \ldots, i_d) \in I$ the probability

$$p_{i_1 \ldots i_{d-1}} \frac{a_i^{(p)}}{\sum_{j=1} a_j^{(p)}}.$$

In other words, we have

$$\tilde{\mu}_p(C_\omega(n)) = \prod_{l=1}^n p_{i_1 \ldots i_{d-1}} \frac{a_i^{(p)}}{\sum_{j=1} a_j^{(p)}}.$$

Let $\mu_p$ be the probability measure on $\Lambda$ which is the pushforward of $\tilde{\mu}_p$ by $\chi$, i.e. $\mu_p = \tilde{\mu}_p \circ \chi^{-1}$.

For calculating the Hausdorff dimension of $\mu_p$, we shall consider some special sets called approximate cubes. Given $\omega \in \Omega$ and $n \in \mathbb{N}$ such that $n \geq (\log \min a_{i_1 \ldots i_d})/(\log \max a_{i_1 \ldots i_d})$, define $L_n^d(\omega) = n$,

$$L_n^d(\omega) = \max \left\{ k \geq 1 : \prod_{l=1}^n a_i^{i_l} \leq \prod_{l=1}^k a_i^{i_l} \right\}$$

$$\vdots$$

$$L_n^{d-1}(\omega) = \max \left\{ k \geq 1 : \prod_{l=1}^n a_i^{i_l} \leq \prod_{l=1}^k a_i^{i_l} \right\}$$

and the approximate cube

$$B_n(\omega) = \left\{ \omega \in \Omega : \frac{i_l}{n} = i_l, \; l = 1, \ldots, n \quad \begin{array}{l} \vdots \\vdots \\vdots \end{array} \right\}$$

We have that each approximate cube $B_n(\omega)$ is a finite union of cylinder sets, and that approximate cubes are nested, i.e., given two, say $B_n(\omega)$ and $B_n'(\omega')$, either $B_n(\omega) \cap B_n'(\omega') = \emptyset$ or $B_n(\omega) \subset B_n'(\omega')$ or $B_n'(\omega') \subset B_n(\omega)$. Moreover, $\chi(B_n(\omega)) = B_n(\omega) \cap \Lambda$ where $B_n(\omega)$ is a closed box in $\mathbb{R}^d$ with edges parallel to the coordinate axes, the $k^{th}$-edge with length $\prod_{l=1}^{L_n^{k-1}(\omega)} a_i^{i_l}$. By (2),

$$1 \leq \frac{\prod_{l=1}^n a_i^{i_l}}{\sum_{l=1}^n a_i^{i_l}} \leq \max a_i^{-1}$$

for $k = 1, \ldots, d$, hence the term “approximate cube”. It follows from (3) that

$$\frac{\sum_{l=1}^{L_n^{k-1}(\omega)} \log a_i^{i_l}}{\sum_{l=1}^n \log a_i^{i_l}} = 1 + \frac{1}{n} \sum_{l=1}^{L_n^{k-1}(\omega)} \log a_i^{i_l} - \frac{1}{n} \sum_{l=1}^n \log a_i^{i_l} \to 1.$$
Also observe that $L_n^{k+1}(\omega) \leq L_n^k(\omega)$ and $L_n^k(\omega) \to \infty$ as $n \to \infty$.

First we calculate the dimension of the “vertical” part. Let
$$J = \{(i^1, \ldots, i^{d-1}) : (i^1, \ldots, i^{d-1}, i^d) \in I \text{ for some } i^d\}$$
and $\Gamma = J^N$. Consider the natural projections $\hat{\pi} : \Omega \to \Gamma$ and $\pi : \mathbb{R}^d \to \mathbb{R}^{d-1}$ given by $\pi(x_1, \ldots, x_d) = (x_2, \ldots, x_d)$. We consider the measures $\hat{\nu}_p = \hat{\mu}_p \circ \hat{\pi}^{-1}$ and $\nu_p = \mu_p \circ \pi^{-1}$.

**Lemma 1.** If $d > 2$ then for every $k \in \{1, \ldots, d-2\}$,
$$\frac{L_n^k(\omega)}{L_n^{k-1}(\omega)} \to \frac{\sum_{i^1 \cdots i^k} p_{i^1 \cdots i^k} \log a_{i^1 \cdots i^k}}{\sum_{i^1 \cdots i^{k+1}} p_{i^1 \cdots i^{k+1}} \log a_{i^1 \cdots i^{k+1}}} \text{ for } \hat{\nu}_p\text{-a.e. } \omega. \tag{5}$$

**Proof.** It follows from (4) that
$$\frac{\sum_{i^1 \cdots i^k} L_n^k(\omega) \log a_{i^1 \cdots i^k}}{\sum_{i^1 \cdots i^{k+1}} L_n^{k+1}(\omega) \log a_{i^1 \cdots i^{k+1}}} \to 1. \tag{5}$$

By Kolmogorov’s Strong Law of Large Numbers (KSLLN),
$$\frac{1}{L_n^{k-1}(\omega)} \sum_{i=1}^{L_n^{k-1}(\omega)} \log a_{i^1 \cdots i^k} \to \sum_{i^1 \cdots i^k} p_{i^1 \cdots i^k} \log a_{i^1 \cdots i^k} \text{ for } \hat{\nu}_p\text{-a.e. } \omega, \tag{6}$$
and (redundantly)
$$\frac{1}{L_n^k(\omega)} \sum_{i=1}^{L_n^k(\omega)} \log a_{i^1 \cdots i^{k+1}} \to \sum_{i^1 \cdots i^{k+1}} p_{i^1 \cdots i^{k+1}} \log a_{i^1 \cdots i^{k+1}} \text{ for } \hat{\nu}_p\text{-a.e. } \omega. \tag{7}$$
The result follows by (5), (6) and (7). \hfill \square

The next lemma is a multidimensional version of [10] Proposition 3.3.

**Lemma 2.** $\dim_H \nu_p = \lambda(p)$. 

**Proof.** To calculate the Hausdorff dimension of $\nu_p$ we are going calculate its pointwise dimension and use [21] Theorem 7.1. Remember that $\chi(B_n(\omega)) = \hat{B}_n(\omega) \cap \Lambda$ where, by [3], $\hat{B}_n(\omega)$ is “approximately” a ball in $\mathbb{R}^d$ with radius $\prod_{i=1}^d a_i^n$, and that
$$\nu_p(\pi \hat{B}_n(\omega)) = \hat{\nu}_p(\hat{B}_n(\omega)).$$

Also, $\chi$ is at most $2^d$ to 1. Taking this into account, by [21] Theorem 7.1| together with [21] Theorem 15.3], one is left to prove that
$$\lim_{n \to \infty} \frac{\log \hat{\nu}_p(\pi \hat{B}_n(\omega))}{\sum_{i=1}^{L_n^d(\omega)} \log a_i} = \lambda(p) \text{ for } \hat{\nu}_p\text{-a.e. } \omega.$$ 

It follows from the definition of $\hat{\nu}_p$ that, for $\hat{\nu}_p$-a.e $\omega$, $p_{i_1 \cdots i_l} > 0$ for every $l$, so we may restrict our attention to these $\omega$. If $d = 2$ then $\hat{\nu}_p(\pi \hat{B}_n(\omega)) = \prod_{i=1}^d a_i^n$ and the result follows by a direct application of KSLLN. Otherwise we have that
$$\hat{\nu}_p(\pi \hat{B}_n(\omega)) = \prod_{i=1}^{d-2} \frac{L_n^{d-2}(\omega)}{L_n^{d-1}(\omega)} \cdot \frac{L_n^{d-1}(\omega)}{L_n^{k+1}(\omega)} \prod_{i=1}^{d-2} \frac{L_n^{d-1}(\omega)}{L_n^d(\omega)}$$
Lemma 3. and this gives what we want after a simple rearrangement.

By Lemma 2, we only have to prove that

\[
\lim_{n \to \infty} \sum_{i=1}^{n} \log a_{i}^{1} = t(p) \quad \text{for } \tilde{\mu}_{p}\text{-a.e. } \omega.
\]

Proof. As before, one is left to prove that

\[
\lim_{n \to \infty} \sum_{i=1}^{n} \log p_{i}^{1} \rightarrow \lambda(p) + t(p) \quad \text{for } \tilde{\mu}_{p}\text{-a.e. } \omega.
\]

We have that

\[
\tilde{\mu}_{p}(B_{n}(\omega)) = \tilde{\nu}_{p}(\bar{\pi}B_{n}(\omega)) \prod_{l=1}^{L_{n}^{d-2}(\omega)} \frac{a_{l}^{1}}{\sum_{i=1}^{d-1} a_{i}^{1}}.
\]

By Lemma 2, we only have to prove that

\[
\lim_{n \to \infty} \sum_{i=1}^{n} \log a_{i}^{1} = t(p) \quad \text{for } \tilde{\mu}_{p}\text{-a.e. } \omega.
\]
But
\[ \frac{\log \alpha_n}{\sum_{i=1}^{\infty} \log a_{ij}} = t(p) \frac{\sum_{i=1}^{d-1} \log a_{ij}^{i_1...i_{d-1}}}{\sum_{i=1}^{d} \log a_{ij}^{i_1...i_{d-1}}} = t(p) \beta_n = \frac{\gamma_n}{\delta_n}. \]

That \( \beta_n \to 1 \) follows from (1). Now we can write
\[ \gamma_n = \sum_{i_1,...,i_{d-1}} P(\omega, L_n^{-1}(\omega), l^{1...d-1}) \log \left( \sum_{i' \in \mathbb{I}} a_{i_1...i_{d-1}}^{i} \right), \]

where
\[ P(\omega, n, l^{1...d-1}) = \sharp \{ 1 \leq l \leq n : (l^{1...i_{d-1}} = (i^{1...i_{d-1}}) \} \]

for \((i^{1,...,i_{d-1}}) \in J\). By KSLLN,
\[ \lim_{n \to \infty} P(\omega, n, l^{1...i_{d-1}}) = p_{i^{1,...,i_{d-1}}} \text{ for } \tilde{\mu}_p\text{-a.e. } \omega, \]

so, by the definition of \( t(p) \),
\[ \gamma_n \to 0 \text{ for } \tilde{\mu}_p\text{-a.e. } \omega. \]

Since \( n / L_n^{-1} (\omega) \geq 1 \), we have that \( |\delta| \geq \log (\min a_{ij}^{-1}) > 0 \), so we also have that
\[ \frac{\gamma_n}{\delta_n} \to 0 \text{ for } \tilde{\mu}_p\text{-a.e. } \omega, \]

thus completing the proof.

As noticed in the beginning of this section, these lemmas imply
\[ \dim_H \Lambda \geq \sup_p \{ \lambda(p) + t(p) \}. \]

**Remark 2.** Theorem 1 can be extended to non-linear self-affine sponges by using the bounded distortion techniques employed in [15] (see also [19]).

4. **A 2-parameter family of Bernoulli measures**

In this section \( d = 3 \) and \( \Lambda \) is a self-affine sponge.

We start by generalizing the definitions of \( t \) and \( t \) for any self-affine sponge. Let \( t = \min_p t(p) \) and \( t = \max_p t(p) \). For \((i,j) \in J\), define \( t_{ij} \) to be the unique real in \([0,1]\) satisfying
\[ \sum_{k=1}^{m_{ij}} a_{ij,k}^{t_{ij}} = 1. \]

It is easy to see that
\[ \ell = \min_{(i,j) \in J} t_{ij} \quad \text{and} \quad \ell = \max_{(i,j) \in J} t_{ij} \]

We observe that the condition \( \ell < \ell \) is open in the numbers \( a_{ij,k} \), so if it is satisfied for an \((a,b,c;0)\)-sponge then it is also satisfied for an \((a,b,c;\varepsilon)\)-sponge, for some \( \varepsilon = \varepsilon(a,I) > 0 \).

Let
\[ P = \left\{ p = (p_{ij})_{(i,j) \in J} : p_{ij} > 0 \text{ for all } (i,j) \in J \text{ and } \sum_{i,j} p_{ij} = 1 \right\}. \]
We will need the following \textit{generic} hypothesis on the numbers \(a_{ijk}\). For each \(t \in (\bar{t}, \bar{t})\), there exist \(1 \leq i \leq m\) and \(1 \leq j < j' \leq m_i\) such that
\[
\sum_{k=1}^{m_{ij}} a_{ij,k}^t \neq \sum_{k=1}^{m_{ij'}} a_{ij',k}^t.
\] (9)

In the next lemma there will be no restrictions on the numbers \(a_{ijk}\) (beside hypothesis (9)). In this way, we say that a self-affine sponge is an \((c, b, \varepsilon)\)-sponge, for some numbers \(0 < b \leq c < 1\), \(\varepsilon \geq 0\), if
\[
eq \frac{b}{b} \leq \varepsilon, \quad \varepsilon \leq \frac{c}{c} \leq \varepsilon,
\]
for every \((i, j) \in J\).

\textbf{Lemma 4.} Let \(0 < b \leq c \leq 1\), \(0 < a_{ij} \leq 1\), \((i, j) \in J\) and \(\delta > 0\). There exists \(\varepsilon = \varepsilon(c, b, a_{ij}, J, \delta)\) such that if \(\Lambda\) is an \((b, c, \varepsilon)\)-sponge satisfying (9) and \(t < \bar{t}\) then:
\[
given t \in [\bar{t} + \delta, \bar{t} - \delta] \text{ and } \rho \in [\delta, 1], \text{ there exists a probability vector } \mathbf{p} = \mathbf{p}(t, \rho) \text{ such that } t(\mathbf{p}) = t \text{ and }
\]
\[
p_{ij} = c_i^{\lambda_i(p)} b_{ij}(p) \left( \sum_k a_{ij,k}^t \right)^{\alpha} \left( \sum_j b_{ij}(p) \left( \sum_k a_{ij,k}^t \right)^{\alpha} \right)^{\rho - 1}, \quad (i, j) \in J,
\]
where \(\alpha = \alpha(t, p) \in \mathbb{R}\) is a \(C^1\) function defined in \([\bar{t} + \delta, \bar{t} - \delta] \times [\delta, 1]\). Moreover, for each \(\rho \in (0, 1]\), \(\alpha(t + \delta, \rho) \to -\infty\) and \(\alpha(t - \delta, \rho) \to \infty\) when \(\delta \to 0\).

\textbf{Proof.} Given \(\alpha, \lambda_1 \in \mathbb{R}\), \(t \in (\bar{t}, \bar{t})\), \(\rho \in (0, 1]\) and \(\lambda_2 \in [0, 1]\), we define a probability vector \(\mathbf{p}(\alpha, \lambda_1, \lambda_2, t, \rho)\) by
\[
p_{ij}(\alpha, \lambda_1, \lambda_2, t, \rho) = C(\alpha, \lambda_1, \lambda_2, t, \rho) c_i^{\lambda_i} b_{ij}(\sum_k a_{ij,k}^t)^{\alpha} \gamma_i(\alpha, \lambda_2, t)^{\rho - 1}
\]
where
\[
\gamma_i(\alpha, \lambda_2, t) = \sum_j b_{ij}(\sum_k a_{ij,k}^t)^{\alpha}
\]
and
\[
C(\alpha, \lambda_1, \lambda_2, t, \rho) = \left( \sum_i c_i^{\lambda_i} \gamma_i(\alpha, \lambda_2, t)^{\rho} \right)^{-1},
\]
for each \((i, j) \in J\).

Let \(F\) be the continuous function defined by
\[
F(\alpha, \lambda_1, \lambda_2, t, \rho) = \sum_{i,j} p_{ij}(\alpha, \lambda_1, \lambda_2, t, \rho) \log(\sum_k a_{ij,k}^t).
\]
We are going to prove there exists a unique \(\alpha = \alpha(\lambda_1, \lambda_2, t, \rho)\), continuously varying, such that \(F(\alpha, \lambda_1, \lambda_2, t, \rho) = 0\), i.e. \(t(\mathbf{p}(\alpha, \lambda_1, \lambda_2, t, \rho)) = t\).

\textbf{Unicity.} We have that, for each \((i, j) \in J\),
\[
\frac{\partial p_{ij}}{\partial \alpha} = \frac{1}{C} \frac{\partial C}{\partial \alpha} p_{ij} \log\left( \sum_k a_{ij,k}^t \right) + p_{ij} + (\rho - 1) \frac{1}{\gamma_i} \frac{\partial \gamma_i}{\partial \alpha} p_{ij}.
\]
Also,
\[
\frac{1}{C} \frac{\partial C}{\partial \alpha} = -\rho \sum_i \frac{p_i}{\gamma_i} \frac{1}{\gamma_i} \frac{\partial \gamma_i}{\partial \alpha}
\] (10)
and
\[
\frac{1}{\gamma_i} \frac{\partial \gamma_i}{\partial \alpha} = \sum_j \frac{p_{ij}}{p_i} \log\left( \sum_k a_{ij,k}^t \right)
\] (11)
where
\[
p_i = \sum_j p_{ij} = C c_i^{\lambda_i} \gamma_i^\rho.
\] (12)
So, by simple rearrangement we get
\[
\frac{\partial F}{\partial \alpha} = \sum_{i,j} \frac{\partial p_{ij}}{\partial \alpha} \log \left( \sum_k a_{ijk}^t \right) = \rho \left\{ \sum p_i \left( \sum_j \frac{p_{ij}}{p_i} \log \left( \sum_k a_{ijk}^t \right) \right)^2 - \left( \sum p_{ij} \log \left( \sum_k a_{ijk}^t \right) \right)^2 \right\} + \sum p_i \left\{ \sum_j \frac{p_{ij}}{p_i} \log \left( \sum_k a_{ijk}^t \right) \right)^2 - \left( \sum_j p_{ij} \log \left( \sum_k a_{ijk}^t \right) \right)^2 \right\}.
\]

By the Cauchy-Schwarz inequality we have that the expressions between curly brackets are non-negative and the second one is positive if there exists \(i \in \{1, \ldots, m\}\) such that the function \(j \mapsto \sum_k a_{ijk}^t \) is non-constant (note that \(p \in \mathcal{P}\)). This is guaranteed by hypothesis (9). Thus \(\frac{\partial F}{\partial \alpha} > 0\).

**Existence.** For fixed \((\lambda_1, \lambda_2, t, \rho)\), we will look at the limit distributions of \(p(\alpha) = p(\alpha, \lambda_1, \lambda_2, t, \rho)\) as \(\alpha \to \pm \infty\).

For \((i, j) \in \mathcal{J}\) such that \(t < t_{ij}\) (remember the definition (8)) we have that
\[
\sum_k a_{ijk}^t > \sum_k a_{ijk}^{t_{ij}} = 1,
\]
so \(A_t = \max_{(i, j) \in \mathcal{J}} \sum_k a_{ijk}^t\).

Consider \((\tilde{i}, \tilde{j}), (i, j) \in \mathcal{J}\) such that
\[
\sum_k a_{ijk}^t < \sum_k a_{ijk}^{t_{ij}} = A_t.
\]

We have
\[
\frac{p_{ij}(\alpha)}{p_{ij}^{(\alpha)}} \leq D \left( \frac{\sum_k a_{ijk}^t}{\sum_k a_{ijk}^{t_{ij}}} \right)^\alpha \left( \frac{\gamma_i(\alpha)}{\gamma_i(\alpha)} \right)^{1-\rho},
\]
for some constant \(D\) not depending on \(\alpha\). Now, for \(\alpha > 0\),
\[
\frac{\gamma_i(\alpha)}{\gamma_i(\alpha)} \leq \tilde{D} \left( \frac{\sum_k a_{ijk}^t}{\sum_k a_{ijk}^{t_{ij}}} \right)^\alpha,
\]
for some constant \(\tilde{D}\) not depending on \(\alpha\). So,
\[
\frac{p_{ij}(\alpha)}{p_{ij}^{(\alpha)}} \leq D\tilde{D} \left( \frac{\sum_k a_{ijk}^t}{\sum_k a_{ijk}^{t_{ij}}} \right)^{\alpha\rho},
\]
which converges to 0 as \(\alpha \to \infty\). This implies that
\[
\sum_{i,j} p_{ij}(\alpha) \log \left( \sum_k a_{ijk}^t \right) \xrightarrow{\alpha \to \infty} \log A_t > 0. \tag{13}
\]

In the same way, defining
\[
B_t = \min_{(i, j) \in \mathcal{J}} \sum_k a_{ijk}^{t_{ij}} < 1,
\]
and taking \((\tilde{i}, \tilde{j}), (i, j) \in \mathcal{J}\) such that
\[
B_t = \sum_k a_{ijk}^{t_{ij}} < \sum_k a_{ijk}^t,
\]
we get, for $\alpha < 0$,
\[ \frac{p_{ij}(\alpha)}{p_{ij}(\alpha)} \leq D \tilde{D} \left( \frac{\sum_k a_{ijk}}{\sum_k a_{ijk}} \right)^{\alpha \rho}, \]
which converges to 0 as $\alpha \to -\infty$. This implies that
\[ \sum_{i,j} p_{ij}(\alpha) \log \left( \sum_k a_{ijk} \right) \xrightarrow{\alpha \to -\infty} \log B_t < 0. \] (14)

By (13), (14) and continuity, there exists $\alpha \in \mathbb{R}$ such that $F(\alpha, \lambda_1, \lambda_2, t, \rho) = 0$. The continuity of $\alpha(\lambda_1, \lambda_2, t, \rho)$ follows from the uniqueness part and the implicit function theorem. Actually, since $F(\alpha, \lambda_1, \lambda_2, t, \rho)$ is continuously differentiable, we also get that $\alpha(\lambda_1, \lambda_2, t, \rho)$ is continuously differentiable. Observe that $t(p) = \tilde{t} \Rightarrow p \in \partial P$ (in this lemma we are assuming $t < \tilde{t}$), so since $t(p(\alpha(\lambda_1, \lambda_2, t, \rho)))) \to \tilde{t}$ when $t \to \tilde{t}$ then
\[ p(\alpha(\lambda_1, \lambda_2, t, \rho)) \to \partial P \quad \text{when} \quad t \to \tilde{t}, \]
which implies
\[ \alpha(\lambda_1, \lambda_2, t, \rho) \to \infty \quad \text{when} \quad t \to \tilde{t}, \]
(this convergence is uniform in $\lambda_1, \lambda_2 \in [0, 1]$, and $\rho$ in a compact set set of $(0, 1)$). In the same way we see that
\[ \alpha(\lambda_1, \lambda_2, t, \rho) \to -\infty \quad \text{when} \quad t \to \underline{t}. \]

Moreover,
\[ \frac{\partial \alpha}{\partial \lambda_1} = - \left( \frac{\partial F}{\partial \alpha} \right)^{-1} \frac{\partial F}{\partial \lambda_1}, \]
where
\[ \frac{\partial F}{\partial \lambda_1} = \sum_{i,j} \frac{\partial p_{ij}}{\partial \lambda_1} \log \left( \sum_k a_{ijk} \right) = \sum_{i,j} p_{ij} \log c_i \log \left( \sum_k a_{ijk} \right), \] (16)
and
\[ \frac{\partial \alpha}{\partial \lambda_2} = - \left( \frac{\partial F}{\partial \alpha} \right)^{-1} \frac{\partial F}{\partial \lambda_2}, \]
where
\[ \frac{\partial F}{\partial \lambda_2} = \sum_{i,j} \frac{\partial p_{ij}}{\partial \lambda_2} \log \left( \sum_k a_{ijk} \right) = \sum_{i,j} p_{ij} \left( \log b_{ij} + (\rho - 1) \frac{1}{\gamma_i} \frac{\partial \gamma_i}{\partial \lambda_2} \right) \log \left( \sum_k a_{ijk} \right), \] (18)
and
\[ \frac{1}{\gamma_i} \frac{\partial \gamma_i}{\partial \lambda_2} = \sum_j \frac{p_{ij}}{p_i} \log b_{ij}. \]

Now we want to find $\lambda_1 = \lambda_1(\lambda_2, t, \rho)$ continuously differentiable such that
\[ C(\alpha(\lambda_1, \lambda_2, t, \rho), \lambda_1, \lambda_2, t, \rho) = 1. \] (20)

We have
\[ \frac{\partial}{\partial \lambda_1} C(\alpha(\lambda_1, \lambda_2, t, \rho), \lambda_1, \lambda_2, t, \rho) = \frac{\partial C}{\partial \alpha} \frac{\partial \alpha}{\partial \lambda_1} + \frac{\partial C}{\partial \lambda_1}, \]
Observe that, by (10) and (11),
\[ \frac{\partial C}{\partial \alpha} = 0 \]
(at points $(\alpha(\lambda_1, \lambda_2, t, \rho), \lambda_1, \lambda_2, t, \rho)$), and
\[ \frac{\partial \log C}{\partial \lambda_1} = \frac{1}{C} \frac{\partial C}{\partial \lambda_1} = \frac{\log C}{\partial \lambda_1} = - \sum_i p_i \log c_i \geq \min_i \log c_i^{-1} > 0. \] (21)
So, \( C(\alpha(\lambda_1, \lambda_2, t, \rho), \lambda_1, \lambda_2, t, \rho) \) is a continuously differentiable function, for each \((\lambda_2, t, \rho)\), is strictly increasing in \(\lambda_1\) and, by (21), has limit \(\infty\) as \(\lambda_1 \to \infty\) and limit 0 as \(\lambda_1 \to -\infty\). By the implicit function theorem, there is a unique \(\lambda_1 = \lambda_1(\lambda_2, t, \rho)\), which is continuously differentiable, satisfying (20). Moreover,

\[
\frac{\partial \lambda_1}{\partial \lambda_2} = - \left( \frac{\partial C}{\partial \lambda_1} \right)^{-1} \frac{\partial C}{\partial \lambda_2} \quad (22)
\]

and

\[
\frac{1}{C} \frac{\partial C}{\partial \lambda_2} = -\rho \sum_i p_i \frac{1}{\gamma_i} \frac{\partial \gamma_i}{\partial \lambda_2} \quad (23)
\]

So, by (21), (23) and (19), we get

\[
\frac{\partial \lambda_1}{\partial \lambda_2} = -\frac{\sum_{i,j} p_{ij} \log b_{ij}}{\sum_i p_i \log c_i} \quad (24)
\]

We see that

\[
\lambda_1(p) = \lambda_1 + \rho \sum_i p_i \log \gamma_i, \quad \lambda_2(p) = \lambda_2 - \sum_{i,j} p_{ij} \log b_{ij}.
\]

We use the following notation

\[
\tilde{\Theta}(\lambda_2, t, \rho) = (\alpha(\lambda_1(\lambda_2, t, \rho), \lambda_2, t, \rho), \lambda_1(\lambda_2, t, \rho), \lambda_2, t, \rho), \quad \Theta(\lambda_2, t, \rho) = (\alpha(\lambda_1(\lambda_2, t, \rho), \lambda_2, t, \rho), \lambda_2, t, \rho).
\]

We want to prove there exists a unique \(\lambda_2 = \lambda_2(t, \rho)\), continuously differentiable, such that

\[
H(\lambda_2, t, \rho) := \sum_i p_i(\tilde{\Theta}) \log \gamma_i(\Theta) = 0.
\]

By (12) and (20),

\[
\frac{\partial}{\partial \lambda_2} \sum_i p_i(\tilde{\Theta}) \log \gamma_i(\Theta) = \frac{\partial \lambda_1}{\partial \lambda_2} \sum_i p_i(\tilde{\Theta}) \log c_i \log \gamma_i(\Theta)
\]

\[
+ \rho \sum_i p_i(\tilde{\Theta}) \frac{1}{\gamma_i(\Theta)} \frac{\partial}{\partial \lambda_2} \gamma_i(\Theta) \log \gamma_i(\Theta) + \sum_i p_i(\tilde{\Theta}) \frac{1}{\gamma_i(\Theta)} \frac{\partial}{\partial \lambda_2} \gamma_i(\Theta)
\]

We have that

\[
\frac{1}{\gamma_i(\Theta)} \frac{\partial}{\partial \lambda_2} \gamma_i(\Theta) = \frac{1}{\gamma_i(\Theta)} \frac{\partial \gamma_i}{\partial \alpha}(\Theta) \left( \frac{\partial \alpha}{\partial \lambda_1} \frac{\partial \lambda_1}{\partial \lambda_2} + \frac{\partial \alpha}{\partial \lambda_2} \frac{\partial \lambda_2}{\partial \lambda_2} \right) + \frac{1}{\gamma_i(\Theta)} \frac{\partial \gamma_i}{\partial \lambda_2}(\Theta).
\]

Then, using (24) and (19), we get

\[
\frac{\partial}{\partial \lambda_2} \sum_i p_i(\tilde{\Theta}) \log \gamma_i(\Theta) = \sum_{i,j} p_{ij}(\tilde{\Theta}) \log b_{ij}
\]

\[
+ \rho \sum_i p_i(\tilde{\Theta}) \left( \sum_{i,j} \frac{p_{ij}(\tilde{\Theta})}{p_i(\tilde{\Theta})} \log b_{ij} - \log c_i \sum_{i,j} p_{ij}(\tilde{\Theta}) \log b_{ij} \right) \log \gamma_i(\Theta)
\]

\[
+ \rho \left( \frac{\partial \alpha}{\partial \lambda_1} \frac{\partial \lambda_1}{\partial \lambda_2} + \frac{\partial \alpha}{\partial \lambda_2} \frac{\partial \lambda_2}{\partial \lambda_2} \right) \sum_i p_i(\tilde{\Theta}) \frac{1}{\gamma_i(\Theta)} \frac{\partial \gamma_i}{\partial \alpha}(\Theta) \log \gamma_i(\Theta).
\]

The term (25) can be made arbitrarily small if \(c_i\) and \(b_{ij}\) are sufficiently close to \(c\) and \(b\), respectively. Note that \(\log \gamma_i(\Theta)\) is uniformly bounded for \(t \in [\bar{t} + \delta, \bar{t} - \delta]\) and \(\rho \in [\delta, 1]\). The term (26) can also be made arbitrarily small if \(c_i\) and \(b_{ij}\) are
sufficiently close to $c$ and $b$, respectively, because, by (15), $\frac{\partial \alpha}{\partial x_1}$ and $\frac{\partial \alpha}{\partial x_2}$ satisfy this property (and the other quantities are uniformly bounded for $t \in [\ell + \delta, \ell - \delta]$ and $\rho \in [\delta, 1]$, see (1)). Then

$$\frac{\partial H}{\partial \lambda_2} = \frac{\partial}{\partial \lambda_2} \sum_i p_i(\Theta) \log \gamma_i(\Theta) < 0$$

if $\Lambda$ is a $(c, b; \varepsilon)$-sponge, for some $\varepsilon = \varepsilon(c, b, a_{ij}, I, \delta) > 0$. Then the existence of $\lambda_2(t, \rho)$ as claimed follows from the implicit function theorem. □

5. Proof of Theorem 2 and Corollary 1

It follows from definitions (2) that, for $(a, b, c; \varepsilon)$-sponges,

$$-\frac{1}{n} + \frac{L_n^1}{n \log b} + \frac{\varepsilon}{\log b} + \frac{\log c}{\log b} \leq \frac{L_n^1}{n} \leq \frac{\log c}{\log b} - \frac{\varepsilon}{\log b} - \frac{L_n^1}{n \log b},$$

and similarly,

$$-\frac{1}{n} + \frac{L_n^2}{n \log a} + \frac{\varepsilon}{\log a} + \frac{\log c}{\log a} \leq \frac{L_n^2}{n} \leq \frac{\log c}{\log a} - \frac{\varepsilon}{\log a} - \frac{L_n^2}{n \log a}.$$

And so,

$$-\frac{1}{n} - A \varepsilon + \frac{\log c}{\log b} \leq \frac{L_n^1}{n} \leq \frac{\log c}{\log b} + A \varepsilon, \quad (27)$$

and

$$-\frac{1}{n} - A \varepsilon + \frac{\log c}{\log a} \leq \frac{L_n^2}{n} \leq \frac{\log c}{\log a} + A \varepsilon, \quad (28)$$

where $A$ is a positive constant (depending only on $a, b, c$) and for every $\varepsilon > 0$ sufficiently small.

We begin by proving the lower estimate in Theorem 2. We leave to the reader to prove that, for $(a, b, c; \varepsilon)$-sponges,

$$\sum_{i=1}^m p_i \log p_i + \sum_{i=1}^m \sum_{j=1}^{m_i} p_{ij} \log p_{ij} - \sum_{i=1}^m p_i \log p_i \geq \sum_{i=1}^m p_i \log p_i + \sum_{j=1}^{m_i} p_{ij} \log b_{ij} \log c$$

where $B$ is a positive constant (depending only on $a, b, c$ and $I$) and for every $\varepsilon > 0$ sufficiently small. Let $t_0$ be such that

$$\sum_{i=1}^m \sum_{j=1}^{m_i} a_{ij}^t = 0.$$

We want to see that $t(p) \geq t_0 - D \varepsilon$ for some positive constant $D$ (depending only on $a, b, c$ and $I$) and for every $\varepsilon > 0$ sufficiently small. This is true if

$$\sum_{i=1}^m \sum_{j=1}^{m_i} p_{ij} \log \sum_{k=1}^{m_{ij}} a_{ijk}^{-t} \geq 0.$$

Now

$$\sum_{i=1}^m \sum_{j=1}^{m_i} p_{ij} \log \sum_{k=1}^{m_{ij}} a_{ijk}^{-t} \geq \sum_{i=1}^m \sum_{j=1}^{m_i} p_{ij} \log \sum_{k=1}^{m_{ij}} (a_{ij}^{-t} a_{ijk}^{-t}) \geq D \varepsilon \log a^{-1} - \varepsilon (t_0 - D \varepsilon) \geq 0,$$

for some appropriate $D$ as before and for every $\varepsilon > 0$ sufficiently small. Then we have that $VP(\Lambda_{a,b,c,\varepsilon}) \geq VP(\Lambda_{a,b,c}) - C \varepsilon$, for some positive constant $C$ (depending
Lemma 5. Let
\[ \frac{\log \mu_p(B_n(\omega))}{\sum_{i=1}^{n} \log c_i} \leq \text{VP}(\Lambda) + C \varepsilon. \]

Remark 3. The fact that \( \dim_H \Lambda_{a,b,c} = \text{VP}(\Lambda_{a,b,c}) \) was essentially proved in [11]. Although in [11] the numbers \( a^{-1}, b^{-1}, c^{-1} \) are assumed to be integers, the proofs work the same when these numbers are not integers. Alternatively, this paper (see the next two lemmas) with \( \varepsilon = 0 \) also gives this result.

Now we prove the upper estimate in Theorem 2.

Lemma 5. Let \( 0 < a \leq b \leq c \leq 1 \) and assume \( t < 1 \). There exists a positive constant \( C \) (depending only on \( a, b, c \) and \( I \)) such that, if \( \epsilon > 0 \) sufficiently small we have the following: if \( \Lambda \) is a \((a,b,c,\varepsilon)\)-sponge and \( \omega \in \Omega \) there exists \( p \in \mathcal{P} \) such that

\[
\liminf_{n \to \infty} \frac{\log \mu_p(B_n(\omega))}{\sum_{i=1}^{n} \log c_i} \leq \text{VP}(\Lambda) + C \varepsilon.
\]

Proof. First assume that \( \Lambda \) satisfy hypothesis [2] (at the end of the proof we say how to deal with the general case). Fix \( \omega \in \Omega \). We use the notation

\[
d_{p,n}(\omega) = \frac{\log \mu_p(B_n(\omega))}{\sum_{i=1}^{n} \log c_i},
\]

Then it follows from the proofs of Lemma 3 and Lemma 2 that, if \( p \in \mathcal{P} \),

\[
d_{p,n}(\omega) = \frac{\sum_{i=1}^{n} \log p_i + \beta_n(\omega) \sum_{i=1}^{n} \log p_{i,j} - \sum_{i=1}^{n} \log p_{i,j}^t}{\sum_{i=1}^{n} \log c_i} + \eta_n(\omega) t(p) = \frac{\sum_{i=1}^{n} \log \left( \sum_k a_{i,j,k}^t(p) \right)}{\sum_{i=1}^{n} \log c_i},
\]

where, by [4],

\[
\beta_n(\omega) = \frac{\sum_{i=1}^{n} \log b_{i,j,k}}{\sum_{i=1}^{n} \log c_i} \to 1 \quad (n \to \infty)
\]

and

\[
\eta_n(\omega) = \frac{\sum_{i=1}^{n} \log a_{i,j,k}^t}{\sum_{i=1}^{n} \log c_i} \to 1 \quad (n \to \infty).
\]

Given \( t \in (0,1) \) and \( \rho \in (0,1] \), consider the probability vector \( p(t,\rho) \), such that \( t(p(t,\rho)) = t \), given by Lemma 4. Applying (29) to \( p(t,\rho) \) we obtain

\[
d_{p(t,\rho),n}(\omega) = \lambda_1(p(t,\rho)) + \beta_n(\omega) \lambda_2(p(t,\rho)) + \eta_n(\omega) t \left( \sum_{i=1}^{n} \log \frac{\gamma_i(t,\rho)}{\sum_{i=1}^{n} \log c_i} \right) + \alpha(t,\rho) \sum_{i=1}^{n} \log \left( \sum_k a_{i,j,k}^t(p) \right),
\]

where

\[
\gamma_i(t,\rho) = \sum_{j=1}^{m_i} b_{i,j} \lambda_2(p(t,\rho)) \left( \sum_{k=1}^{m_j} a_{i,j,k}^t(p) \right)^{\alpha(t,\rho)}.
\]
We choose \( \rho = \frac{\log c}{\log b} \) and \( t \) such that \( \alpha(t, \rho) = \frac{\log b}{\log c} \). Using estimates (27), (28) and the fact that we are considering an \((a, b, c; \varepsilon)\)-sponge, we get
\[
d_{p(t, \rho), n}(\omega) \leq \lambda_1(p(t, \rho)) + \beta_2(\omega)\lambda_2(p(t, \rho)) + \eta_n(\omega)t
\]
\[
= \frac{\log c}{\log b} \sum_{i=1}^{\log b} \log \left( \sum_{j=1}^{m_{ij}} m_{ij} \right) - \sum_{i=1}^{\log b} \log \left( \sum_{j=1}^{m_{ij}} m_{ij} \right)
\]
\[
+ \frac{\log b}{\log a} \sum_{i=1}^{\log a} \log m_{i} - \sum_{i=1}^{\log a} \log m_{i} + \tilde{c}_\varepsilon + \frac{D}{n}
\]
for some constants \( \tilde{C} \) and \( D \) (depending only on \( a, b, c \) and \( I \)) and for every \( \varepsilon > 0 \) sufficiently small. By [11, Lemma 4.1] we have that
\[
\limsup_{n \to \infty} \frac{1}{n} \left( \frac{\log c}{\log b} \sum_{i=1}^{\log b} \log \left( \sum_{j=1}^{m_{ij}} m_{ij} \right) - \sum_{i=1}^{\log b} \log \left( \sum_{j=1}^{m_{ij}} m_{ij} \right) \right)
\]
\[
+ \frac{\log b}{\log a} \sum_{i=1}^{\log a} \log m_{i} - \sum_{i=1}^{\log a} \log m_{i} \geq 0,
\]
so
\[
\liminf_{n \to \infty} d_{p(t, \rho), n}(\omega) \leq \lambda_1(p(t, \rho)) + \lambda_2(p(t, \rho)) + t + \tilde{C}_\varepsilon \leq \nu \Lambda + \tilde{C}_\varepsilon.
\]

Now we deal with hypothesis (9). Given \( \eta > 0 \), since the quantities in this lemma depend continuously on the numbers \( a_{ijk} \), we can substitute \( a_{ijk} \) by arbitrarily close numbers \( \bar{a}_{ijk} \) that satisfy hypothesis (2) and so that at the end we obtain
\[
\liminf_{n \to \infty} d_{p(t, \rho), n}(\omega) \leq \nu \Lambda + \tilde{C}_\varepsilon + \eta.
\]
Since \( \eta \) can be made arbitrarily close to zero by making \( \bar{a}_{ijk} \) sufficiently close to \( a_{ijk} \), we get the desired result.

**Lemma 6.** Let \( 0 < a < b < c \leq 1 \) and assume \( \xi < \eta \). There exists a positive constant \( \tilde{C} \) (depending only on \( a, b, c \) and \( I \)) such that, for every \( \varepsilon > 0 \) sufficiently small, if \( \Lambda \) is an \((a, b, c; \varepsilon)\)-sponge then
\[
\dim_H \Lambda \leq \nu \Lambda + \tilde{C}_\varepsilon.
\]

**Proof.** Let \( \xi > 0 \). Consider the approximate cubes of order \( n \) given by \( B_n(z) = \chi(B_n(\omega)) \) where \( \omega \in \chi^{-1}(z), z \in \Lambda, n \in \mathbb{N} \). Then it follows from Lemma [5] that
\[
\forall z \in \Lambda \forall N \in \mathbb{N} \exists n > N \exists \mu \in \mathcal{P} : \frac{\log \mu(B_n(z))}{\log |B_n(z)|} \leq \nu \Lambda + \tilde{C}_\varepsilon + \xi.
\]
(30)

Given \( \delta, \eta > 0 \), we shall build a cover \( \mathcal{U}_{\delta, \eta} \) of \( \Lambda \) by sets with diameter \( \eta \) such that
\[
\sum_{U \in \mathcal{U}_{\delta, \eta}} |U|^{\nu \Lambda + \tilde{C}_\varepsilon + \xi + 2\delta} \leq \sqrt{3} (\max a_{ijk}) M_\delta
\]
where \( M_\delta \) is an integer depending on \( \delta \) but not on \( \eta \). This implies that \( \dim_H \Lambda \leq \nu \Lambda + \tilde{C}_\varepsilon + \xi + 2\delta \) which gives what we want because \( \xi \) and \( \delta \) can be taken arbitrarily small. Let \( c = \max c_i < 1 \). It is clear that there exists a finite number of Bernoulli measures \( \mu_1, \ldots, \mu_M_\delta \) such that
\[
\forall \mu \exists k \in \{1, \ldots, M_\delta\} : \frac{\mu(B_n)}{\mu(B_n)} \leq c^{-\delta n}
\]
for all approximate cubes of order \( n \), \( B_n \). By (30), we can build a cover of \( \Lambda \) by approximate cubes \( B_n(z^i), i = 1, 2, \ldots \) that are disjoint and have diameters \( \eta \), such that
\[
\mu(B_n(z^i)) \geq |B_n(z^i)|^{\nu \Lambda + \tilde{C}_\varepsilon + \xi + \delta}
\]
for some probability vectors $\mathbf{p}$. It follows that
\[
\sum_i |B_n(z_i)|^{\mu(p)} + \varepsilon + 2^N \leq \sum_i \mu_{\mathbf{p}}(B_n(z_i)) |B_n(z_i)|^\delta \\
\leq \sum_i \mu_{k_i}(B_n(z_i)) \varepsilon^{-\delta n(z_i)} \sqrt{3} (\max a_{i,j,k}^{-1}) \varepsilon^{\delta n(z_i)} \\
\leq \sqrt{3} (\max a_{i,j,k}^{-1}) \sum_k \sum_i \mu_k(B_n(z_i)) \leq \sqrt{3} (\max a_{i,j,k}^{-1}) M_{\delta}
\]
as we wish. □

The upper estimate of Corollary 1 follows from Lemma 6. As in the beginning of this section, we have that $\text{VP}(\Lambda_{a,b,c}; \varepsilon) \leq \text{VP}(\Lambda_{a,b,c}) + D \varepsilon$ for some positive constant $D$ (depending only on $a, b, c$ and $I$) and for every $\varepsilon > 0$ sufficiently small. This together with Lemma 6 gives the upper estimate in Theorem 2.

REFERENCES

[1] K. Baranski, Hausdorff dimension of the limit sets of some planar geometric constructions, Adv. Math. 210 (2007), 215-245.
[2] J. Barral, D.-J. Feng, Non-uniqueness of ergodic measures with full Hausdorff dimension on Gatzouras-Lalley carpet, Nonlinearity 24 (2011), 2563-2567.
[3] J. Barral, D.-J. Feng, Dimensions of random statistically self-affine Sierpinski esponjes in $\mathbb{R}^k$, J. Math. Pures Appl 149 (2021), 254-303.
[4] T. Bedford, Crinkly curves, Markov partitions and box dimension of self similar sets, PhD Thesis, University of Warwick, 1984.
[5] R. Bowen, Equilibrium states and the ergodic theory of Anosov diffeomorphisms (Lecture Notes in Mathematics, 470). Springer, 1975.
[6] R. Bowen, Hausdorff dimension of quasi-circles, Publ. Math. I.H.E.S. 50 (1979), 11-26.
[7] T. Das and D. Simmons, The Hausdorff and dynamical dimensions of self-affine sponges: a dimension gap result, Invent. Math. 226 (2011), 2470-2502.
[8] D. Gatzouras and P. Lalley, Hausdorff and box dimensions of certain self-affine fractals, Indiana Univ. Math. J. 41 (1992), 533-568.
[9] D. Gatzouras and Y. Peres, Invariant measures of full dimension for some expanding maps, Ergod. Th. & Dynam. Sys. 17 (1997), 147-167.
[10] I. Heuer and S. Lalley, Falconer’s formula for the Hausdorff dimension of a self-affine set in $\mathbb{R}^2$, Ergod. Th. & Dynam. Sys. 15 (1995), 97-97.
[11] M. Hochman, A. Rapaport, Hausdorff dimension of planar self-affine sets and measures with overlaps, J. Eur. Math. Soc. (2021), published online first.
[12] R. Kenyon and Y. Peres, Measures of full dimension on affine-invariant sets, Ergod. Th. and Dynam. Sys. 16 (1996), 307-323.
[13] N. Luzia, A variational principle for the dimension of a class of non-conformal repellors, Ergod. Th. & Dynam. Sys. 26 (2006), 821-845.
[14] N. Luzia, Hausdorff dimension for an open class of repellors in $\mathbb{R}^2$. Nonlinearity 19 (2006), 2895-2908.
[15] N. Luzia, Measure of full dimension for a class of nonconformal repellors, Discrete Contin. Dyn. Syst. 26 (2010), 291-302.
[16] N. Luzia, Hausdorff dimension of certain random self-affine fractals, Stoch. Dyn. 11 (2011), 627-642.
[17] N. Luzia, On the uniqueness of an ergodic measure of full dimension for non-conformal repellors, Discrete Contin. Dyn. Syst. 37 (2017), 5763-5780.
[18] C. McMullen, The Hausdorff dimension of general Sierpinski carpets, Nagoya Math. J. 96 (1984), 1-9.
[19] Ya. Pesin, Dimension Theory in Dynamical Systems: Contemporary Views and Applications, Chicago University Press, 1997.
[20] D. Ruelle, Repellers for real analytic maps, Ergod. Th. and Dynam. Sys. 2 (1982), 99-107.
[21] D. Ruelle, Thermodynamic Formalism, Second Edition, Cambridge University Press, 2004.
[24] Ya. Sinai, Gibbs measures in ergodic theory, *Uspehi Mat. Nauk* 27 (1972), 21-64. English translation: *Russian Math. Surveys* 27 (1972), 21-69.

**Universidade Federal do Rio de Janeiro, Instituto de Matemática, Rio de Janeiro 21941-909, Brazil**

*Email address:* nuno@im.ufrj.br