STATIONARY MEASURES FOR PROJECTIVE TRANSFORMATIONS: THE BLACKWELL AND FURSTENBERG MEASURES

B. BÁRÁNY, M. POLLICOTT AND K. SIMON

ABSTRACT. In this paper we study the Blackwell and Furstenberg measures, which play an important role in information theory and the study of Lyapunov exponents. For the Blackwell measure we determine parameter domains of singularity and give upper bounds for the Hausdorff dimension. For the Furstenberg measure, we establish absolute continuity for some parameter values. Our method is to analyze linear fractional iterated function schemes which are contracting on average, have no separation properties (that is, we do not assume that the open set condition holds, see [9]) and, in the case of the Blackwell measure, have place dependent probabilities. In such a general setting, even an effective upper bound on the dimension of the measure is difficult to achieve.

1. INVARIANT MEASURES FOR PROJECTIVE TRANSFORMATIONS

In this note we will be interested in the problem of understanding invariant measures for an interesting class of iterated function schemes of the interval. Such problems have been extensively studied in the case of affine iterated function schemes, and there have been a number of partial results in more general settings. However, in this note we will consider the interesting case of linear fractional transformations, whose significance is that they arise naturally from projective transformations. To avoid too much abstraction, we will concentrate on two cases which have historically been particularly important. Namely, the Blackwell measure and the Furstenberg measure. The fundamental question we want to address in each case is when these measures are absolutely continuous or singular.

The Blackwell measure, at least in the particular case of binary symmetric channels, is an invariant measure for linear fractional transformations with a certain class of rational weights. This measure was introduced by Blackwell in 1957 and plays a central role in understanding the entropy rate, and other important characteristics, of fundamental models in information theory. We show that for a suitable range of parameter values this measure is necessarily singular. In order to study

the regions upon which stationary measures are singular (in particular in the case of the Blackwell measures), we will employ some recent results established by Jaroszewska and Rams [10].
Another important invariant measure for linear fractional iterated function schemes, arising this time in the study of Lyapunov exponents for matrices, is the Furstenberg measure. This too arises from linear fractional transformations, but this time with constant weights. In this case we can show the more positive result that for almost all parameters in a certain range of values the Furstenberg measure is absolutely continuous. We recall that there are results of a complementary nature by Kaimanovich and Le Prince [11] in the case of Fuchsian groups with Bernoulli weights. In that paper, the authors construct an example where the corresponding boundary measure is singular and conjectured that this was the typical situation. Whereas our results are not directly applicable, they do not support their conjecture.
In section 2 we specialise to the case of the Blackwell measure, and in section 3 we concentrate on a particular class of examples corresponding to certain binary symmetric channels with noise. In section 4, we give new criteria for the singularity of the Blackwell measure, and illustrate this with a simple example.
In section 6, we turn to the Furstenberg measure, and in section 7 we concentrate on a particular class of examples corresponding to a pair of positive 2×2 matrices. In section 8 we give criteria for the absolute continuity of the Furstenberg measure, and again illustrate this with a simple example.
In the next section we will consider natural examples of contracting on average IFS with place dependent probabilities and no separation properties. Such examples arise quite naturally in the study of certain practical problems. Indeed, these appear to be the first significant examples of such systems for which dimension theoretical properties are treated. Moreover, in these examples we can make precise numerical estimates on $h(\mu)/\lambda(\mu)$.

2. The Blackwell measure

In this section we recall the general construction of the Blackwell measure [2]. Consider a stationary ergodic Markov process

$$Y := \{y_n\}_{n=-\infty}^{\infty} \text{ with states } i = 1, \ldots, B,$$

with transition matrix $M = [m_{ij}]_{ij=1}^{B}$ and associated probability measure $\mathbb{P}$. Then the entropy $H(Y)$ is given by

$$H(Y) = -\sum_{ij=1}^{B} \mathbb{P}\{y_n = i\} \cdot m(i,j) \cdot \log m(i,j).$$

Assume that $A < B$, then for any surjective map $\Phi\{1,\ldots,B\} \to \{1,\ldots,A\}$ we can consider the ergodic stationary process

$$Z := \{z_n = \Phi(y_n)\}_{n=-\infty}^{\infty} \text{ with states } i = 1,\ldots,A.$$

Blackwell [2] expressed the the entropy for Z as follows:

$$H(Z) = -\int \sum_{a=1}^{A} r_a(\mathbf{w}) \log r_a(\mathbf{w}) dQ(\mathbf{w}), \tag{1}$$

where the measure Q is called the *Blackwell measure* and Q can be characterized as the invariant measure of an Iterated Function Scheme below. (Figure 2 gives a picture of the definitions to be introduced in a special case.) First we define

$$r_a(\mathbf{w}) := \sum_{i=1}^{B} \sum_{\{j:\ \Phi(j)=a\}} w_i \cdot m(i,j) \tag{2}$$

for $a \in \{1,\cdots,A\}$. Note that $(r_1(\mathbf{w}),\ldots,r_A(\mathbf{w}))$ is a probability vector for all

$$\mathbf{w} \in W := \left\{ \mathbf{w} \in \mathbb{R}^B : w_i \geq 0, \sum_{i=1}^{B} w_i = 1 \right\}. \tag{3}$$

To define the Blackwell measure Q we introduce the functions

$$f_a(\mathbf{w}) := \sum_{j=1}^{B} \mathbf{e}_j \cdot \delta_a(\Phi(j)) \cdot \sum_{i=1}^{B} w_i \cdot m(i,j)/r_a(\mathbf{w}), \tag{4}$$

where $\mathbf{e}_j$ is the j-th coordinate unit vector. Note that $f : W \to W_a$, where for $a = 1,\ldots,A$ we define

$$W_a := \{\mathbf{w} \in W : w_j = 0 \text{ if } \Phi(j) \neq a\}.$$

Now we are ready to define Q as the invariant measure of the Iterated Function Scheme:

$$\{f_1(\mathbf{w}),\ldots,f_A(\mathbf{w})\} \text{ with probabilities } \{r_1(\mathbf{w}),\ldots,,r_A(\mathbf{w})\}.$$

That is for a Borel set $E \subset W$ we have

$$Q(E) = \sum_{a=1}^{A} \int_{f_a^{-1}(E)} r_a(\mathbf{w}) dQ(\mathbf{w}).$$

Alternatively, for a continuous function $F : W \to \mathbb{R}$ we have

$$\int F(\mathbf{w}) dQ(\mathbf{w}) = \sum_{a=1}^{A} \int r_a(\mathbf{w}) \cdot F(f_a(\mathbf{w})) dQ(\mathbf{w}).$$

Clearly we have

$$\mathrm{suppQ} \subset \bigcup_{a=1}^{A} W_a.$$

3. A class of examples of the Blackwell measure

In this section we focus on a most frequently studied example of hidden Markov chain (see e.g. [7, Example 4.1]) :

Example 1. *We set $B = 4, A = 2$ and let $\Phi : \{1, 2, 3, 4\} \to \{1, 2\}$ be given by*

$$\Phi(1) = \Phi(4) = 1 \text{ and } \Phi(2) = \Phi(3) = 2.$$

In this case we have

$$W := \left\{ \mathbf{w} = (w_1, \ldots, w_4) \in \mathbb{R}^4 : w_i \geq 0 \text{ and } \sum_{i=1}^{4} w_i = 1 \right\},$$

and

$$W_1 := \{\mathbf{w} \in W : \mathbf{w} = (w_1, 0, 0, w_4)\} \text{ and } W_2 = \{\mathbf{w} \in W : \mathbf{w} = (0, w_2, w_3, 0)\}.$$

In this case the support of the Blackwell measure $\mathrm{supp}(Q) = W_1 \cup W_2$.

The importance of this Example 1 is that it describes the binary channel with corrupted output. Namely, the input of the binary symmetric channel is a binary symmetric Markov chain source $\{X_i\}_{i=-\infty}^{\infty}$, $X_i \in \{0, 1\}$ with probability transition matrix

$$\Pi := \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}.$$

The channel adds to X a binary i.i.d. noise $E = \{E_i\}_{i=-\infty}^{\infty}$ with

$$\mathrm{Prob}(E_i = 0) = 1 - \varepsilon, \qquad \mathrm{Prob}(E_i = 1) = \varepsilon.$$

The corrupted output is the stationary stochastic process $Z = \{Z_i\}_{i=-\infty}^{\infty}$

$$Z_i = X_i \bigoplus E_i,$$

where $\bigoplus$ denotes the binary addition. Clearly, $Y := \{Y_i\}_{i=-\infty}^{\infty}$, $Y_i = (X_i, E_i)$ is a Markov chain with states $(0,0), (0,1), (1,0), (1,1)$ and transition probabilities:

$$M := \begin{bmatrix} p(1-\varepsilon) & p\varepsilon & (1-p)(1-\varepsilon) & (1-p)\varepsilon \\ p(1-\varepsilon) & p\varepsilon & (1-p)(1-\varepsilon) & (1-p)\varepsilon \\ (1-p)(1-\varepsilon) & (1-p)\varepsilon & p(1-\varepsilon) & p\varepsilon \\ (1-p)(1-\varepsilon) & (1-p)\varepsilon & p(1-\varepsilon) & p\varepsilon \end{bmatrix}$$

Theorem 2. *The Blackwell measure is singular in the blue region marked on Figure 1.*

We will prove the theorem later in Section 4, page 17.

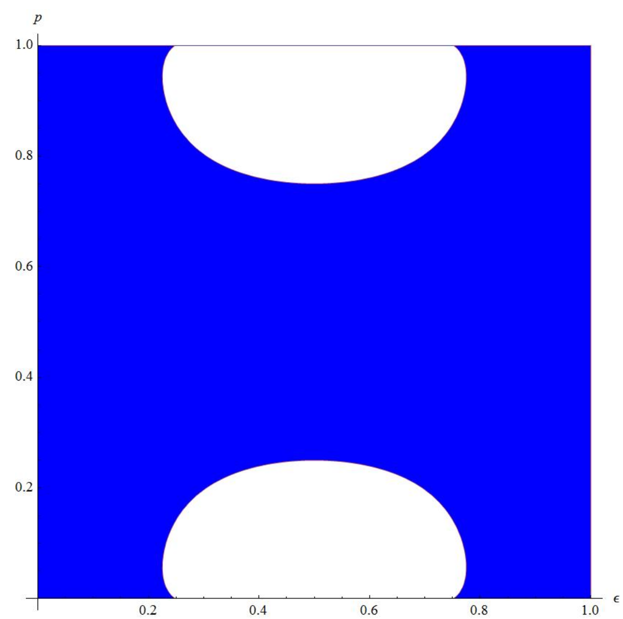

FIGURE 1. Estimates on where the Blackwell measure is singular (via higher iterates).

3.1. **The projected conjugate IFS $\{S_1, S_2\}$ on $[0,1]$.** In this section we introduce an IFS on the unit interval to analyze the entropy rate. This content is not new. It has been established earlier in [8] and [15]. In these cases the Markov input sources were not assumed to be symmetric.

Using the definitions introduced in (2), (3) and (4) the Blackwell measure for this canonical example is the invariant measure of the IFS

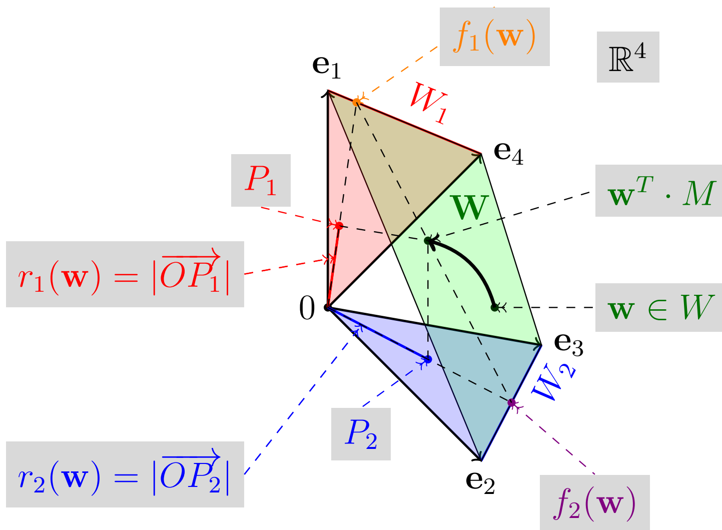

FIGURE 2. W is the simplex spanned by the coordinate unit vectors $\mathbf{e}_1, \ldots, \mathbf{e}_4$. P_1, P_2 are the orthogonal projection of $\mathbf{w}^T M$ to the coordinate plane spanned by $(\mathbf{e}_1, \mathbf{e}_4)$ and $(\mathbf{e}_2, \mathbf{e}_3)$ respectively. $r_i(\mathbf{w}) = \|OP_i\|_1$, $f_i(\mathbf{w})$ is the intersection of the lines OP_i and W_i.

$\{f_1, f_2\}$ acting on W with place dependent probabilities $\{p_1, p_2\}$. See Figure 2 for the geometric interpretation of these functions.
To get the formulaes for $r_a(\mathbf{w}), f_a(\mathbf{w})$ we introduce

$$M_1 := \begin{bmatrix} p(1-\varepsilon) & 0 & 0 & (1-p)\varepsilon \\ p(1-\varepsilon) & 0 & 0 & (1-p)\varepsilon \\ (1-p)(1-\varepsilon) & 0 & 0 & p\varepsilon \\ (1-p)(1-\varepsilon) & 0 & 0 & p\varepsilon \end{bmatrix}, M_2 := \begin{bmatrix} 0 & p\varepsilon & (1-p)(1-\varepsilon) & 0 \\ 0 & p\varepsilon & (1-p)(1-\varepsilon) & 0 \\ 0 & (1-p)\varepsilon & p(1-\varepsilon) & 0 \\ 0 & (1-p)\varepsilon & p(1-\varepsilon) & 0 \end{bmatrix}.$$

Then for $a = 1, 2$ we have that

$$r_a(\mathbf{w}) = \mathbf{w} \cdot M_a \cdot \mathbf{1} =: \|w.M_a\|_1 \text{ and } f_a(\mathbf{w}) := \mathbf{w} \cdot M_a / r_a(\mathbf{w}).$$

Note that the first two and the last two rows of the matrix M (and so of the matrices M_1, M_2) are identical. This implies that for the most natural parameterizations of W_1, W_2,

$$\mathbf{w}_1(t) := (t, 0, 0, 1-t), \mathbf{w}_2(t) := (0, t, 1-t, 0), \ t \in [0, 1]$$

we have

$$\mathbf{w}_1(t) \cdot M_a = \mathbf{w}_2(t) \cdot M_a \quad a = 1, 2.$$

So, by definition,

$$f_a(\mathbf{w}_1(t)) = f_a(\mathbf{w}_2(t)) \text{ and } r_a(\mathbf{w}_1(t)) = r_a(\mathbf{w}_2(t)) \quad a = 1, 2.$$

So, we can define the place dependent probabilities $\{p_1(t), p_2(t)\}$ and the IFS $\{S_1(t), S_2(t)\}$:

$$p_a(t) := r_a(\mathbf{w}_b(t)) \text{ and } S_a(t) := \mathbf{w}_a^{-1}\left(f_a(\mathbf{w}_b(t))\right), a = 1,2,\ b = 1,2.$$

That is for $f_{a,b} := f_a|_{W_b}$. the following diagram commutes:

$$\begin{array}{ccc} [0,1] & \xrightarrow{S_a} & [0,1] \\ \mathbf{w}_b \downarrow & & \downarrow \mathbf{w}_a \\ W_b & \xrightarrow[f_{a,b}]{} & W_a \end{array} \tag{5}$$

This shows the reason that we introduced the IFS $\{S_1, S_2\}$.

$$p_1(t) = t \cdot [p(1-\varepsilon) + (1-p)\cdot\varepsilon] + (1-t)\cdot[(1-p)(1-\varepsilon) + p\cdot\varepsilon]$$

and

$$p_2(t) = t \cdot [p\varepsilon + (1-p)\cdot(1-\varepsilon)] + (1-t)\cdot[(1-p)\varepsilon + p\cdot(1-\varepsilon)].$$

Further,

$$S_1^{\varepsilon,p} := S_1(t) := [t\cdot p\cdot(1-\varepsilon) + (1-t)\cdot(1-p)\cdot(1-\varepsilon)]/p_1(t) \tag{6}$$

and

$$S_2^{\varepsilon,p} := S_2(t) := [t\cdot p\cdot\varepsilon + (1-t)\cdot(1-p)\cdot\varepsilon]/p_2(t).$$

To further simplify the calculations we introduce

$$q := \varepsilon(1-p) + p(1-\varepsilon).$$

It is easy to see that if $0 < \varepsilon, p \le \frac{1}{2}$ or $\frac{1}{2} \le \varepsilon, p \le 1$ then $0 < q \le \frac{1}{2}$. If from ε, p only one of them is greater than $1/2$ then we can apply that $S_1^{\varepsilon,p} \equiv S_2^{1-\varepsilon,p}$, $S_2^{\varepsilon,p} \equiv S_1^{1-\varepsilon,p}$ and $p_1^{\varepsilon,p} \equiv p_2^{1-\varepsilon,p}$, $p_2^{\varepsilon,p} \equiv p_1^{1-\varepsilon,p}$ which puts us to the situation considered above. Therefore without loss of generality we may assume that $0 < q < \frac{1}{2}$. Immediate calculation yields that

Lemma 3. *We can write*

$$\begin{aligned} p_1(x) &= (2q-1)x + 1 - q, \\ p_2(x) &= (1-2q)x + q. \end{aligned} \tag{7}$$

In this way $p_1(x)$ is always monotone decreasing and $p_2(x)$ is always monotone increasing. In particular we have that

$$\begin{aligned} p_1(0) &= p_2(1) = 1 - q, \\ p_2(0) &= p_1(1) = q. \end{aligned}$$

Moreover, for every $x \in [0,1]$ we have the identity

$$S_1(x) + S_2(1-x) = 1.$$

Clearly, $p_1(x) + p_2(x) \equiv 1$. In the sequel we frequently use the notation

$$\overleftarrow{\mathbf{i}} := (i_n, \dots, i_1) \text{ if } \mathbf{i} = (i_1, \dots, i_n).$$

Further,

$$p_{\mathbf{i}}(x) := p_x(\mathbf{i}) := p_{i_1}(x) p_{i_2}(S_{i_1}(x)) \cdots p_{i_n}(S_{i_{n-1} \dots i_1}(x)). \tag{8}$$

Fact 4.

- **(a):** $p_x(\mathbf{i})$ *is a probability measure on* Σ_2 *for every* $x \in [0,1]$.
- **(b):** $p_x(\mathbf{ij}) = p_x(\mathbf{i}) \cdot p_{S_{\overleftarrow{\mathbf{i}}}(x)}(\mathbf{j})$
- **(c):** *There exist* $C_1, C_2 > 0$ *such that for every* n*, for every* $\mathbf{i} \in \{1,2\}^n$ *and for all* $x, y \in [0,1]$ *we have*

$$C_1 < \frac{p_{\mathbf{i}}(x)}{p_{\mathbf{i}}(y)} < C_2.$$

The proofs of (a) and (b) are obvious, while the proof of (c) follows from the next Lemma which states that our IFS is eventually contracting.

Lemma 5. *There exists an* n *such that for every* $(i_1, \dots, i_n) \in \{1,2\}^n$ *and for every* $x \in [0,1]$ *we have*

$$\left| S'_{i_1 \dots i_n}(x) \right| < 1.$$

Proof. An immediate calculation yields that the compact set $J := S_1([0,1]) \cup S_2([0,1])$ is contained in the open interval $(0,1)$. Let $\widetilde{J}_k := w_k(J) \subset W_k$ for $k = 1,2$. Then $\widetilde{J}_k$ is a compact subset of the interior of W_k. Let

$$L := \max\left\{ |S'_i(x)| : x \in [0,1] \text{ and } i = 1,2 \right\}.$$

Then by [7, Proposition 2.2] we can choose an n such that for every $(i_1, \dots, i_{n-1}) \in \{1,2\}^{n-1}$ the map $f_{i_1 \dots i_{n-1}}|_{\widetilde{J}_1 \cup \widetilde{J}_2}$ is a contraction with ratio less than $1/2L$.
By the chain rule, for an arbitrary $x \in [0,1]$ we have

$$S'_{i_1 \dots i_n}(x) = S'_{i_1 \dots i_{n-1}}(S_{i_n}(x)) \cdot S'_{i_n}(x).$$

The commutative diagram (5) yields that

$$\begin{aligned} S_{i_1 \dots i_{n-1}}(y) &= w_{i_1}^{-1} \circ f_{i_1} \circ w_{i_2} \circ \cdots w_{i_{n-2}}^{-1} \circ f_{i_{n-2}} \circ w_{i_{n-1}} w_{i_{n-1}}^{-1} \circ f_{i_{n-1}} \circ w_{i_n} \\ &= w_{i_1}^{-1} \circ f_{i_1 \dots i_{n-1}} \circ w_{i_n}(y). \end{aligned}$$

Since both $w_1(x)$ and $w_2(x)$ are similarities with ratio $\sqrt{2}$, the map $S_{i_1\dots i_{n-1}}|_J$ is a contraction with ration less than $1/2L$. Thus we obtain that for every $(i_1,\dots,i_n)\in\{1,2\}^n$ and $x\in[0,1]$,

$$|S'_{i_1\dots i_n}(x)| < \frac{1}{2}.$$

□

An easy calculation now shows the following:

Lemma 6. *We can write that*

$$\begin{aligned} S_1'(x) &= \frac{\Delta}{p_1^2(x)}, \text{ and} \\ S_2'(x) &= \frac{\Delta}{p_2^2(x)}, \end{aligned} \tag{9}$$

where

$$\Delta := 2\varepsilon p - 2\varepsilon^2 p - \varepsilon + \varepsilon^2 = -\varepsilon(1-\varepsilon)(1-2p).$$

This then gives the following corollary:

Corollary 7.

1. *For $p<1/2$ both $S_1(x)$ and $S_2(x)$ are monotone decreasing.*
2. *For $p>1/2$ both $S_1(x)$ and $S_2(x)$ are monotone increasing.*
3. *For $p=1/2$ both of the functions $S_1(x)$ and $S_2(x)$ are constant.*

So, in what follows we always assume that

$$p \neq \frac{1}{2}.$$

3.2. **The projections of the Blackwell measure to the line.** We know that the Blackwell measure Q is supported by $W_1\cup W_2$. Let us write $Q_a := Q|_{W_a}$ for $a=1,2$. This leads naturally to the definition of the Borel measures m_1, m_2 and m on the interval $[0,1]$ defined by

$$m_a := \left(\mathbf{w}_a^{-1}\right)_* Q_a, \text{ for } a=1,2 \text{ and } m(B) := m_1(B)+m_2(B). \tag{10}$$

That is

$$m_a(B) := Q_a(\mathbf{w}_a(B)) \text{ for } a=1,2.$$

In this way, for a Borel function $G: W_1\cup W_2\to\mathbb{R}$

$$\int_0^1 (G\circ\mathbf{w}_a)(t)dm_a(t) = \int_{W_a} G(\mathbf{w})dQ_a(\mathbf{w}).$$

We apply this for $G(\mathbf{w}) = \sum_{b=1}^{2} r_b(\mathbf{w}) \log r_b(\mathbf{w})$ and we use that $r_b(\mathbf{w}_a(t)) = p_b(t)$ for all $a = 1, 2$, $b = 1, 2$ to obtain that

$$\int_{W_a} G(\mathbf{w}) dQ_a(\mathbf{w}) = \int_0^1 \sum_{b=1}^{2} p_b(t) \log p_b(t) dm_a(t).$$

Hence, by (1) the entropy of the hidden Markov chain is

(11)

$$H(Z) = -\sum_{a=1}^{2} \int_{W_a} G(\mathbf{w}) dQ_a(\mathbf{w}) = -\int_0^1 p_1(t) \log p_1(t) + p_2(t) \log p_2(t) dm(t).$$

Using (10) and (5) we obtain from the change of variable formula that for a Borel set $B_1 \subset [0, 1]$ and $E_1 = \mathbf{w}_1(B_1) \subset W_1$:

$$\begin{aligned} m_1(B_1) &= Q_1(E_1) = \int_{f_1^{-1}(E_1)} r_1(\mathbf{w}) dQ(\mathbf{w}) \\ &= \int_{f_{1,1}^{-1}(E_1)} r_1(\mathbf{w}) dQ_1(\mathbf{w}) + \int_{f_{1,2}^{-1}(E_1)} r_1(\mathbf{w}) dQ_2(\mathbf{w}) \\ &= \int_{S_1^{-1}(B_1)} p_1(x) dm_1(x) + \int_{S_1^{-1}(B_1)} p_1(x) dm_2(x) \\ &= \int_{S_1^{-1}(B_1)} p_1(x) dm(x). \end{aligned}$$

Similarly, for $E_2 = \mathbf{w}_2(B_2) \subset W_2$:

$$m_2(B_2) = \int_{f_2^{-1}(E_2)} r_2(\mathbf{w}) dQ(\mathbf{w}) = \int_{S_2^{-1}(B_2)} p_2(x) dm(x).$$

Adding the last two formulaes together we obtain that

$$m(B) = \int_{S_1^{-1}(B)} p_1(x) dm(x) + \int_{S_2^{-1}(B)} p_2(x) dm(x). \tag{12}$$

Thus it follows that for every n we have that $\mathrm{supp}(m) \subset \cup_{i_1 \cdots i_n} S_{i_1 \cdots i_n}[0, 1]$. Therefore, if $\max |S_1'(x)| + \max |S_2'(x)| < 1$ then $m \perp \mathrm{leb}$. It is immediate from the definition of the Hausdorff dimension (see [5]) that

$$\dim_{\mathrm{H}}(Q) = \dim_{\mathrm{H}}(m),$$

so it is enough to estimate the left hand side. For the definition and basic properties of the Hausdorff dimension. Let us call $\mathcal{S}$ the IFS which consists of the functions $\{S_1(x), S_2(x)\}$, chosen with the place dependent probabilities $\{p_1(x), p_2(x)\}$, respectively. It follows from

Lemma 5 that $\mathcal{S}$ is eventually contracting. By (12) the measure m is an invariant measure for the IFS $\mathcal{S}$.

We fix an N which satisfies Lemma 5. Then the IFS

$$\widehat{\mathcal{S}} := \{S_{\mathbf{i}}\}_{\mathbf{i}\in\{1,2\}^N}$$

with probabilities $p_{\mathbf{i}}(x)$ (defined in (8)) is strictly contracting. We write $\widehat{T}$ for the N-th iterate of the operator $T : \mathcal{C}^1([0,1]) \to \mathcal{C}^1([0,1])$

$$(Tf)(x) := p_{i_1}(x) \cdot f(S_{i_1}(x)) + p_{i_2}(x) \cdot f(S_{i_2}(x)). \tag{13}$$

That is for an $f \in \mathcal{C}^1([0,1])$ we have

$$(\widehat{T}f)(x) = \sum_{\mathbf{i}\in\{1,2\}^N} p_{\overleftarrow{\mathbf{i}}}(x) \cdot f(S_{\mathbf{i}}(x)).$$

Using that the IFS $\widehat{\mathcal{S}}$ is contracting, a standard and easy calculation shows that there exists a constant $C > 0$ such that for any $f \in \mathcal{C}^1([a,b])$ and $n \in \mathbb{N}$,

$$\|(\widehat{T}^n f)'\|_\infty \le C\|f\|_\infty + \|f'\|_\infty. \tag{14}$$

The following lemma summarizes two of the important properties of the measure m.

Lemma 8. *The measure m is the unique invariant measure for the IFS $\mathcal{S}$. Moreover, it is necessarily of pure type (i.e., it is either singular or absolutely continuous with respect to the Lebesgue measure).*

Proof. Let f be a C^1 function, then we claim that $T^n f$ converges to a constant in the C^1 topology. Observe that it follows from (14) that the sequence of $\mathcal{C}^1[0,1]$ functions $\{\widehat{T}^n f\}_{n=1}^\infty$ is an equicontinuous family. Thus by the Azela-Ascoli theorem we can deduce that there is a convergent subsequence $\{T^{n_i} f\}_{i=1}^\infty$ to a continuous function $\overline{f}$, say. Furthermore,

$$\inf_x f(x) \le \inf_x Tf(x) \le \inf_x T^2 f(x) \le \cdots \le \inf_x \overline{f}(x).$$

and so for every n we have $\inf_x \overline{f}(x) = \inf_x T^n \overline{f}(x)$. Moreover, since for any x the set $\{S_{i_1} \cdots S_{i_n}(x) : i_1, \cdots, i_n, n \ge 1\}$ is dense, we can deduce that $\overline{f}(x) \equiv \inf_x \overline{f}(x)$, i.e., the limit is a constant $c(f)$, say. Furthermore, $T^n f \to c(f)$ as $n \to +\infty$ and if $T_*\mu = \mu$ is an invariant measure then $\mu(f) = \mu(T^n f) \to c(f)$. In particular, since $c(f)$ is independent of μ we deduce that there is a unique T-invariant measure. The uniqueness of the T-invariant measure μ leads easily to the conclusion that μ is of pure type. If we write $\mu = \mu_{ac} + \mu_{sing}$ for the unique

Lebesgue decomposition into absolutely continuous and singular parts then we observe that $T_*\mu_{ac} = \mu_{ac}$ and $T_*\mu_{sing} = \mu_{sing}$. However, by the uniqueness of the stationary measure we can deduce that one of these vanishes.

□

Remark 9. *For any $g \in \mathcal{C}^1([0,1])$ we have that $T^n g \to \int g d\mu(x)$ uniformly (at an exponential rate). In practice the rate of convergence might be difficult to compute. However, some rigorous estimates can come either using the Birkhoff cone method* [17] *or cycle expansions* [4].

3.3. **The measure m is the push down measure of a Gibbs measure μ.** Here we give an important characterization of the measure m. Let $\Sigma := \{1,2\}^{\mathbb{N}}$ and $\varphi : \Sigma \to \mathbb{R}$ be defined by

$$\varphi(\mathbf{i}) := \log p_{i_1}(\Pi(\sigma\mathbf{i})),$$

where, as usual, $\Pi : \Sigma \to [0,1]$ denotes the natural projection. That is

$$\Pi(\mathbf{i}) := \lim_{n\to\infty} S_{i_1\dots i_n}(0).$$

Clearly, φ is Hölder continuous. So, it follows from [1, Theorem 1.4] that there is a Gibbs measure μ on Σ for the potential φ. Then μ is a σ (left shift on Σ) invariant probability measure. This means that for every $f \in \mathcal{C}(\Sigma)$ we have

$$\begin{aligned}\int_\Sigma f(\mathbf{j})d\mu(\mathbf{j}) &= \int_\Sigma \sum_{k=1}^{2} \exp\left(\varphi(k\mathbf{j})\right) f(k\mathbf{j})d\mu(\mathbf{j}) \\ &= \int_\Sigma \big(\underbrace{p_1(\Pi\mathbf{j})f(1\mathbf{j}) + p_2(\Pi\mathbf{j})f(2\mathbf{j})}_{(\mathcal{L}f)(\mathbf{j})}\big)d\mu(\mathbf{j}). \qquad (15)\end{aligned}$$

As usual we write $\Pi_*\mu$ for the push down measure of μ. That is $\Pi_*\mu(H) := \mu(\Pi^{-1}(H))$.
Using that $p_1(x) + p_2(x) \equiv 1$ we get that for the transfer operator $\mathcal{L} : \mathcal{C}(\Sigma) \to \mathcal{C}(\Sigma)$,

$$(\mathcal{L}f)(\mathbf{j}) := p_1(\Pi\mathbf{j})f(1\mathbf{j}) + p_2(\Pi\mathbf{j})f(2\mathbf{j})$$

we have $\mathcal{L}(1) = 1$. That is the constant function $h(\mathbf{j}) \equiv 1$ is an eigenfunction corresponding to the maximal eigenvalue $\lambda = 1$. This implies that the topological pressure $P = 0$ since it was proved in [1, p. 26] that $\lambda = \mathrm{e}^P$.

Remark 10. *We define a projection from $\widetilde{\Pi} : C[0,1] \to C(\Sigma)$ as follows*

$$\left(\widetilde{\Pi} g\right)(\mathbf{i}) := g(\Pi(\mathbf{i})).$$

Then an immediate calculation yields that for every $g \in C(\Sigma)$ and $\mathbf{i} \in \Sigma$ we have

$$\left[\left(\widetilde{\Pi} \circ T\right)(g)\right](\mathbf{i}) = \left[\left(\mathcal{L}\widetilde{\Pi}\right)(g)\right](\mathbf{i})$$

that is the following diagram commutes:

$$\begin{array}{ccc} C[0,1] & \xrightarrow{T} & C[0,1] \\ \widetilde{\Pi}\downarrow & & \downarrow\widetilde{\Pi} \\ C(\Sigma) & \xrightarrow[\mathcal{L}]{} & C(\Sigma) \end{array}$$

Now we prove that m is the push down measure of the Gibbs measure μ.

Proposition 11. *We have*

$$m = \Pi_* \mu.$$

Proof. Let $g : [0,1] \to \mathbb{R}$ be an arbitrary continuous function. Then by the change of variables formulae

$$\mathcal{I} := \int_0^1 g(x) d(\Pi_*\mu)(x) = \int_\Sigma (g \circ \Pi)(\mathbf{j}) d\mu(\mathbf{j}).$$

Using this and (15) for $f = g \circ \Pi$ we obtain

$$\mathcal{I} = \int_\Sigma \left(p_1(\Pi\mathbf{j}) \cdot g\left(S_1(\Pi\mathbf{j})\right) + p_2(\Pi\mathbf{j}) \cdot g\left(S_2(\Pi\mathbf{j})\right)\right) d\mu(\mathbf{j}).$$

Applying the change of variables formulae again we get

$$\int_0^1 g(x) d(\Pi_*\mu)(x) = \int_0^1 (Tg)(x) d(\Pi_*\mu)(x).$$

That is $T^*(\Pi_*\mu) = \Pi_*\mu$. Since there is a unique invariant probability measure for T, we have $m = \Pi_*\mu$. □

3.4. **The entropy and Lyapunov exponent.** Since $m = \Pi_*\mu$, as usual, we define the entropy of m as

$$h(m) := h(\mu).$$

We note that using (12) it is possible to define the entropy of m as done in [10]. However, it is much more convenient for us to define the entropy $h(m)$ as $h(\mu)$.

Using that μ is the Gibbs measure for the potential φ and the variational principle (see [1, Theorem 1.2]) we get:

$$0 = P(\varphi) = h(\mu) + \int_{\Sigma} \varphi d\mu.$$

That is applying the definition of φ, formulae (15) and the change of variables formulae in this order we get

$$\begin{aligned} h(m) = h(\mu) &= -\int_{\Sigma} \log p_{i_1}(\Pi(\sigma \mathbf{i})) d\mu(\mathbf{i}) \\ &= -\int_{\Sigma} \left(p_1(\Pi\mathbf{i}) \log p_1(\Pi\mathbf{i}) + p_2(\Pi\mathbf{i}) \log p_2(\Pi\mathbf{i})\right) d\mu(\mathbf{i}) \qquad (16) \\ &= -\int_0^1 \left(p_1(x) \log p_1(x) + p_2(x) \log p_2(x)\right) dm(x). \end{aligned}$$

Put

$$\mathfrak{h}(x) := -\left(p_1(x) \log p_1(x) + p_2(x) \log p_2(x)\right)$$

By (11) and (16) we get that the entropy $H(Z)$ of the hidden Markov chain under consideration and the entropy of m are equal to the integral of $\mathfrak{h}(x)$:

$$H(Z) = h(m) = h(\mu) = \int_0^1 \mathfrak{h}(x) dm(x). \qquad (17)$$

As usual [13] we define the Lyapunov exponent of the measure m by

$$\lambda(m) := \int_{\Sigma} \log |S'_{i_1}(\Pi(\sigma \mathbf{i}))| d\mu(\mathbf{i}).$$

To express it in a more convenient way, we introduce the function

$$\ell(x) = p_1(x) \log |S_1'(x)| + p_2(x) \log |S_2'(x)|.$$

Then like in (16) one can easily see that

$$\lambda(m) = \int_0^1 \ell(x) dm(x). \qquad (18)$$

An immediate consequence of Lemma 6 is the following corollary:

Corollary 12. $\ell(x) = \log |\Delta| + 2\mathfrak{h}(x)$.

Hence the Lyapunov exponent defined in our paper is equal to $\log|\Delta| + 2h(m)$. As we mentioned earlier, the ratio entropy / Lyaponov exponent is used to estimate the Hausdorff dimension from above.

Corollary 13. *We have that*

$$\frac{h(m)}{-\lambda(m)} \le a \le 1$$

if and only if

$$H(Z) = h(m) = \int \mathfrak{h}(x) dm(x) \le -\frac{a}{1+2a} \cdot \log|\Delta|.$$

4. Singularity of the Blackwell measure

We are now in a position to study ranges of parameter values for which the Blackwell measure is singular. The dimension of a measure is defined as the infimum of the Hausdorff dimension of full measure sets (see [13]). So, if the dimension of a measure is less than 1 then the measure must be singular w.r.t. the Lebesgue measure. Using this principle, we find large (ε, p) parameter areas where the corresponding Blackwell measure is singular because of the ratio entropy/Lyapunov exponent is less than 1 for the measure m.

Proposition 14.

$$\dim_{\mathrm{H}}(m) \le \frac{h(m)}{-\lambda(m)}.$$

The proof of the Proposition follows from the proof of [13, Theorem 4.4.2]. Namely, in [13, Theorem 4.4.2] the proof of the upper bound does not use any separation conditions like open set condition, see [9]. Alternatively, the assertion of this Proposition readily follows from the much more general [10, Theorem 1].

It follows from (18), Corollary 12 and (17) that we get a better understanding of the the ratio entropy/Lyapunov exponent by analyzing the function $\mathfrak{h}(x)$. Some elementary calculation yields that:

Lemma 15. *The function $\mathfrak{h}(x)$ satisfies:*

- **(a):** $0 < \mathfrak{h}(x) = \mathfrak{h}(1-x)$ *for all* $x \in [0,1]$,
- **(b):** $\mathfrak{h}''(x) \le 0$ *for all* $x \in [0,1]$,
- **(c):** $\mathfrak{h}(x)$ *attains its maximum at* $x = \frac{1}{2}$,
- **(d):** $\mathfrak{h}(0) = \mathfrak{h}(1)$ *is the value of the minimum of the function* $\mathfrak{h}(x)$.

The standard proof of Lemma 15 can be found for example in [3].

Definition 16. *We say the a smooth real function f defined on $[0,1]$ is a* reversed cup function *if f satisfies the prosperities (a)-(d) above.*

Lemma 17. *Let $f(x)$ be a reversed cup function. Then the function $(Tf)(x)$ is also a reversed cup function.*

Proof. Using (9) and the chain rule we obtain that

$$\begin{aligned}(Tf)'(x) &= (2q-1)\cdot f(S_1(x)) + (1-2q)\cdot f(S_2(x)) \\ &+ f'(S_1(x))\cdot \underbrace{S_1'(x)\cdot p_1(x)}_{\Delta/p_1(x)} \\ &+ f'(S_2(x))\cdot \underbrace{S_2'(x)\cdot p_2(x)}_{\Delta/p_2(x)}\end{aligned}$$

Similarly, using that $\Delta, S_1'(x), S_2'(x)$ have the same sign we get

$$\begin{aligned}(Tf)''(x) &= (2q-1)\cdot f'(S_1(x))\cdot S_1'(x) + (1-2q)\cdot f'(S_2(x))\cdot S_2'(x) \\ &+ f''(S_1(x))\cdot S_1'(x)\cdot \frac{\Delta}{p_1(x)} + f''(S_2(x))\cdot S_2'(x)\cdot \frac{\Delta}{p_2(x)} \\ &+ f'(S_1(x))\cdot(-1)\cdot \underbrace{\Delta\cdot p_1^{-2}(x)}_{S_1'(x)}(2q-1) \\ &+ f'(S_2(x))\cdot(-1)\cdot \underbrace{\Delta\cdot p_2^{-2}(x)}_{S_2'(x)}(1-2q) \\ &= f''(S_1(x))\cdot S_1'(x)\cdot \frac{\Delta}{p_1(x)} + f''(S_2(x))\cdot S_2'(x)\cdot \frac{\Delta}{p_2(x)} \qquad (19) \\ &\leq 0.\end{aligned}$$

In this way we have checked that property (b) in Lemma 15 holds. To see that property (a) holds as well we need to observe that

$$\begin{aligned}(Tf)(x) &= p_1(x)f(S_1(x)) + p_2(x)f(S_2(x)) \\ &= p_1(x)f(S_1(x)) + p_1(1-x)f(\underbrace{1-S_1(1-x)}_{S_2(x)}) \\ &= p_1(x)f(S_1(x)) + p_1(1-x)f(S_1(1-x)) \\ &= (Tf)(1-x).\end{aligned}$$

This completes the proof. □

This leads to the next result.

Proposition 18. *For every n we have*

$$(T^n\mathfrak{h})(0) \leq h(m) \leq (T^n\mathfrak{h})(1/2).$$

Proof. Since $T^*m = m$ it follows from (16) that

$$h(m) = \int \mathfrak{h}(x)dm(x) = \int (T^n\mathfrak{h})(x)dm(x).$$

Using Lemma 15 and Lemma 17 we obtain that $(T^n\mathfrak{h})(x)$ is a reversed cup function. This immediately implies the assertion of the Proposition. □

Proof of Theorem 2. It follows form Corollary 13, Proposition 14 and Proposition 18 that for an arbitrary (ε, p), if we can find an n such that

$$3(T^n\mathfrak{h})(1/2) + \log|\Delta| < 0$$

holds then the Blackwell measure corresponding to parameters (ε, p) is singular. On Figure 1 one can see the region of (ε, p) where $3(T^{10}\mathfrak{h})(1/2) + \log|\Delta| < 0$. □

Example 19. *Let us consider a concrete example:* $p = 0.2$ *and* $\epsilon = 0.3$. *By the preceding we can sandwich it in the nested intervals* $[(T^n\mathfrak{h})(1/2)], (T^n\mathfrak{h})(0)]$. *In the following table we present these values for* $n = 1, \cdots, 15$.

n	$(T^n\mathfrak{h})(1/2)$	n	$(T^n\mathfrak{h})(0)$
0	0.6931471805599453094	0	0.6640641265641080113
1	0.6885320764504560426	1	0.6807022640392729670
2	0.6873399079169119219	2	0.6852455729100437792
3	0.6870204663672780007	3	0.6864604838094890865
4	0.6869351099501216909	4	0.6867853827159174941
5	0.6869122847651288517	5	0.6868722510733396336
6	0.6869061819804526502	6	0.6868954778673486178
7	0.6869045502238403004	7	0.6869016881839058585
8	0.6869041139291720920	8	0.6869033486838228601
9	0.6869039972737272948	9	0.6869037926642685923
10	0.6869039660826768519	10	0.6869039113746924475
11	0.6869039577428888054	11	0.6869039431151996655
12	0.6869039555130163633	12	0.6869039516019000626
13	0.6869039549167984387	13	0.6869039538710535796
14	0.6869039547573831256	14	0.6869039544777743832
15	0.6869039547147590431	15	0.6869039546399979308

Thus we can get bounds on the ration h/χ of the blackwell measure in terms of the parameter values (p, ϵ).

Remark 20. *Unfortunately, there don't appear to be any techniques available to show that the Blackwell measure is absolutely continuous on particular domains.*

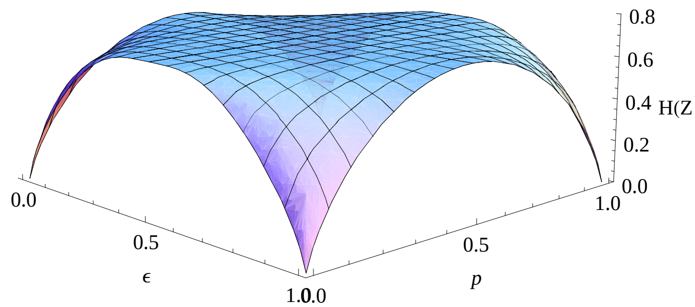

FIGURE 3. The entropy of the Hidden Markov chain as a function of ε, p.

5. ESTIMATE OF THE ENTROPY OF THE HIDDEN MARKOV CHAIN

Our aim in this chapter is to find rigorous error estimate on the computation as we have seen in (16) the entropy of the Hidden Markov chain

$$H(Z) = \lim_{n\to\infty} (T^n \mathfrak{h})(0) = \lim_{n\to\infty} (T^n \mathfrak{h})(\frac{1}{2}).$$

Method 1
We define v such that

$$[v, 1-v] = \text{convex hull}\left(S_1([0,1]) \cup S_2([0,1])\right).$$

It follows from definition (6) that $0 < v < \frac{1}{2}$. All the inequalities we obtain below immediately follow form the fact that f is a reversed cup function and form the definition (13) of the transfer operator T. First note that $(Tf)(\frac{1}{2}) \leq f(\frac{1}{2})$ and there exists a $u \in \left(v, \frac{1}{2}\right)$ such that $f(u) = (Tf)(0)$. Then for some $t_2 \in (v, \frac{1}{2})$ we have

$$(Tf)\left(\frac{1}{2}\right) - (Tf)(0) < f\left(\frac{1}{2}\right) - f(u) \leq f\left(\frac{1}{2}\right) - f(v) = f'(t_2)\left(\frac{1}{2} - v\right)$$

On the other hand, for some $t_1 \in (0, v)$ we have

$$f(v) - f(0) = v \cdot f'(t_1).$$

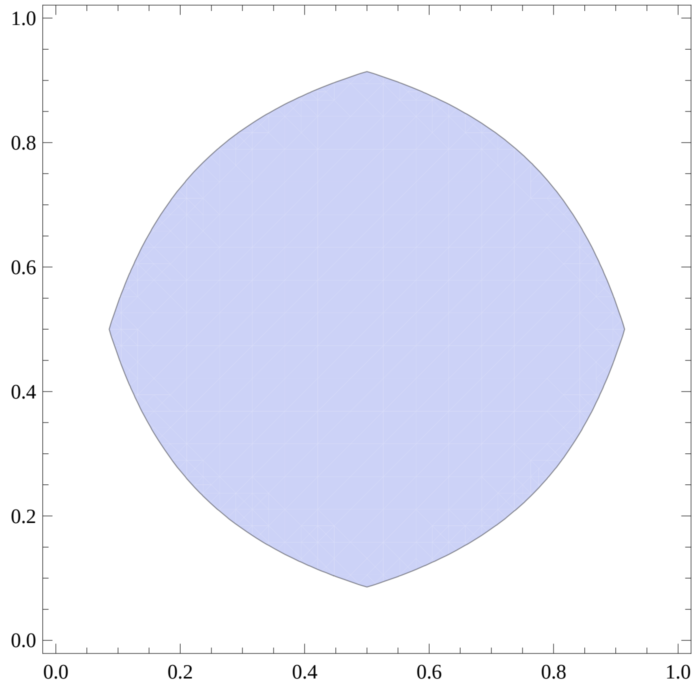

FIGURE 4. The region where the error is less than 0.1 after 14 iteration of the transfer operator using Method 1.

Putting the last two formulaes together and using that $f'(t_1) > f'(t_2) > 0$ we get

$$\begin{aligned}\frac{f\left(\frac{1}{2}\right)-f(0)}{(Tf)\left(\frac{1}{2}\right)-(Tf)(0)} &\geq \frac{f\left(\frac{1}{2}\right)-f(v)}{f\left(\frac{1}{2}\right)-f(v)}+\frac{f(v)-f(0)}{f\left(\frac{1}{2}\right)-f(v)} \\ &= 1+\frac{v\cdot f'(t_1)}{\left(\frac{1}{2}-v\right)\cdot f'(t_2)} \\ &\geq 1+\frac{v}{\frac{1}{2}-v}=\frac{1}{1-2v}\end{aligned}$$

This yields that

$$\frac{(Tf)\left(\frac{1}{2}\right)-(Tf)(0)}{f\left(\frac{1}{2}\right)-f(0)} \leq 1-2v.$$

Note that

$$\mathfrak{h}\left(\frac{1}{2}\right)-\mathfrak{h}(0)=\log 2+((1-q)\log(1-q)+q\log q)\,.$$

Hence

$$|H(Z)-(T^n\mathfrak{h})\,(0)| < (1-2v)^n\cdot(\log 2+(1-q)\log(1-q)+q\log q)\,.$$

As we have seen v, q depends on ε, p.

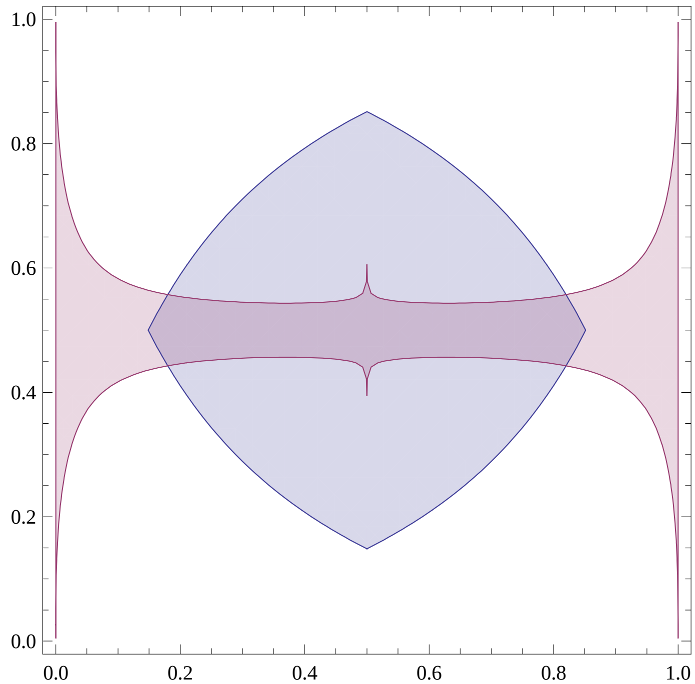

FIGURE 5. The region where the error is less than 0.0001 after 14 iteration of the transfer operator. The rhomboid region comes from Method 1 and the region like H is resulted by Method 2.

Our second method is as follows:
Method 2 Let g be a reversed cup function. By Taylor formulae there exits a $u \in \left(0, \frac{1}{2}\right)$ such that

$$g(0) = g\left(\frac{1}{2}\right) - \frac{1}{2} \cdot g'\left(\frac{1}{2}\right) + \frac{1}{8} \cdot g''(u).$$

Using that $g'(1/2) = 0$ and introducing the notation

$$M_{2,g} := \max_{x \in [0,1]} |g''(x)|$$

we obtain that

$$\left| g(0) - g\left(\frac{1}{2}\right) \right| \leq \frac{1}{8} M_{2,g}. \tag{20}$$

In lights of this, to estimate $\left| (T^n \mathfrak{h})(0) - (T^n \mathfrak{h})\left(\frac{1}{2}\right) \right|$ it is enough to give an upper bound on $M_{2,T^n \mathfrak{h}}$. By (19) and Lemma 6 for a reversed cup function f we get

$$(Tf)''(x) = f''(S_1(x)) \cdot \frac{\Delta}{p_1^3(x)} + f''(S_2(x)) \cdot \frac{\Delta}{p_2^3(x)}.$$

From this and (7) we conclude that

$$M_{2,Tf} \leq M_{2,f} \cdot \frac{|\Delta|}{q^3(1-q)^3}.$$

So, we start with $\mathfrak{h}''(x) = -(1-2q)^2 \cdot \frac{1}{p_1(x)p_2(x)}$. That is

$$M_{2,\mathfrak{h}} = \frac{(1-2q)^2}{q(1-q)}.$$

Put $\gamma := |\Delta|/(1-q)^3 q^3$. Then by the last two displayed equations.

$$M_{2,T^n\mathfrak{h}} \leq \frac{(1-2q)^2}{q(1-q)} \cdot \gamma^n.$$

This and (20) yields

$$|H(Z) - T^n\mathfrak{h}(0)| \leq \left| (T^n\mathfrak{h})(0) - (T^n\mathfrak{h})\left(\frac{1}{2}\right) \right| \leq \frac{1}{8} \cdot \frac{(1-2q)^2}{q(1-q)} \cdot \gamma^n$$

5.1. **Numerical Approximation.** The numerical error is even smaller then it was presented in the previous methods. One can use the program "Mathematica 8" to estimate numerically the error. Using the function `NMaximize` with algorithm `NelderMead` one can show that the the numerical error of the 10th approximation of the entropy is at most 0.026 (for further information about the program and the algorithm see [20]). The upper bound of the error of the 10th approximation of the entropy (see Figure 3), precisely, the graph of the function $(\varepsilon, p) \mapsto T^{10}\mathfrak{h}(1/2) - T^{10}\mathfrak{h}(0)$ can be seen on Figure 6.

6. The Furstenberg measure

We now turn to the complementary problem of understanding the nature of the Furstenberg measure. We begin with the basic definition. The Furstenberg measure is associated to a finite set of matrices $A_1, \cdots .A_k \in GL(d, \mathbb{R})$ chosen randomly with respect to a given iid Bernoulli measure associated to $(p_1, \cdots, p_k)$, say, and is used to characterize the lyapunov exponents.

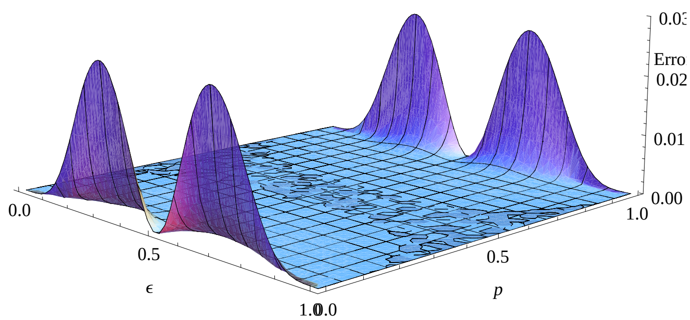

FIGURE 6. The error of the approximation of the entropy.

6.1. The Furstenberg measure.

Definition 21. *The Lyapunov exponent λ is given by the limit*

$$\lambda := \lim_{n\to+\infty} \frac{1}{n} \sum_{i_1,\cdots,i_n} p_{i_1}\cdots p_{i_n} \log ||A_{i_1}\cdots A_{i_n}||$$

By a famous result of Kesten and Furstenberg one has the following pointwise estimate with repesct to the Bernoulli measure $\mu = (p_1, \cdots, p_k)^{\mathbb{Z}^+}$:

Theorem 22 (Furstenberg-Kesten)**.** *For a.e. (μ) $\underline{i} \in \Sigma$ one has*

$$\lambda = \lim_{n\to+\infty} \frac{1}{n} \log ||A_{i_1}\cdots A_{i_n}||.$$

Thus the Lyapunov exponent λ estimates the rate of growth of the norm

$$||A_{i_1}A_{i_2}A_{i_3}\cdots A_{i_n}||$$

for a typical random product of matrices A_{i_1}, A_{i_2}, $A_{i_3}, \ldots$ A_{i_n}, with respect to the Euclidean norm $\|\cdot\|$ on $\mathbb{R}^d$, say.
We now turn to the associated Furstenberg measure, which is defined on projective space $P\mathbb{R}^{d-1}$. Let $(p_1, \cdots, p_k)$ be a probability vector, and let $\mu = (p_1, \cdots, p_k)^{\mathbb{Z}^+}$ be the associated Bernoulli measure on the space of sequences $\Sigma = \{1, \cdots, k\}^{\mathbb{Z}^+}$. More precisely, consider the real projective sphere

$$P\mathbb{R}^{d-1} = (\mathbb{R}^d - \{(0,0)\})/\sim,$$

where $v \sim w$ if there exists $\beta > 0$ such that $\beta v = w$. The usual linear action $A_j : \mathbb{R}^d \to \mathbb{R}^d$ allows one to define a projective action $A_j : P\mathbb{R}^{d-1} \to P\mathbb{R}^{d-1}$ by

$$A_j(x_1, \cdots, x_d) = \frac{A_j x}{\sqrt{\sum_{i=1}^d (A_j x)_i^2}},$$

for $j = 1, \cdots, k$.

Standing assumption: Henceforth, we shall always assume that the matrices are all positive (i.e., each of the entries of each of the matrices $A_1, \cdots, A_k$ are strictly positive).

In particular, we can then deduce that the induced action on projective space is a contraction, with respect to a suitable metric. More precisely, by the well known Birkhoff cone theorem, we know that the associated actions contract with respect to the Hilbert-Birkhoff metric on projective space.
This implies (c.f. [9]) that there is a unique probability measure ν on $P\mathbb{R}^d$ satisfying

$$\nu(E) = \sum_{j=1}^k p_j \nu(A_j^{-1}(E)).$$

Actually we obtain this measure as

$$\nu(E) = \Pi_*(\mu) \text{ where } \Pi(\mathbf{i}) := \lim_{n\to\infty} A_{i_1} \cdots A_{i_n}(w),$$

for an arbitrary $w \in P\mathbb{R}^d$. One of the reasons that this measure ν is important is as follows.

Theorem 23 (Furstenberg [6]). *The* Standing hypothesis *implies that*

$$\lambda = \int_{P\mathbb{R}^d} \sum_{j=1}^k p_j \cdot \log \|A_j(x)\| d\nu(x).$$

Therefore the measure ν is called *Furstenberg measure.*

7. A class of examples of the Furstenberg measure

In this section we consider the particular case that $d = 2$ and $p_1 = p_2 = \frac{1}{2}$. We can associate to a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, say, a linear fractional transformation on the unit interval $[0, 1]$ denoted by

$$g : x \mapsto \frac{ax + b(1-x)}{(a+c)x + (b+d)(1-x)}.$$

In particular, this corresponds to the projective action for a particular choice of chart. Furthermore, if we assume that $d = 2$ we can associate to two such matrices

$$A_0 = \begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix} \text{ and } A_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$$

the corresponding orientation preserving maps $g_0, g_1 : [0,1] \to [0,1]$ are denoted by

$$g_0 : x \mapsto \frac{a_0 x + b_0(1-x)}{(a_0+c_0)x + (b_0+d_0)(1-x)} \text{ and } g_1 : x \mapsto \frac{a_1 x + b_1(1-x)}{(a_1+c_1)x + (b_1+d_1)(1-x)}$$

The hypothesis on the positivity of the matrices makes it natural to make the following simplifying assumptions.

(1) the maps are contracting, i.e., $||g_0'||_\infty, ||g_1'||_\infty < 1$
(2) the maps g_0, g_1 have distinct fixed points $g_0(x_1) = x_0$ and $g_1(x_1) = x_1$.

We can assume without loss of generality that $x_0 < x_1$ then the limit set of the iterated function scheme lies in the interval $[x_0, x_1]$ and so after a simple affine change of coordinate we can assume without loss generality that $x_0 = 0$ and $x_1 = 1$. In particular, in these coordinates we can assume that

$$g_0 : x \mapsto \frac{\alpha_0 x}{x + d_0(1-x)} \text{ and } g_1 : x \mapsto \frac{d_1 x + (1-\alpha_1)(1-x)}{d_1 x + (1-x)},$$

The functions g_0 and g_1 depend on altogether four parameters from which we form the vector $\mathbf{t} := (\alpha_0, \alpha_1, d_0, d_1) \in (\mathbb{R}^+)^4$. As in the general case here we also write $\nu = \nu_{\mathbf{t}}$ for the *Furstenberg measure*. That is ν is the unique probability measure on the interval $[0,1]$ satisfying

$$\nu(E) = \frac{1}{2} \cdot \nu\left(g_0^{-1}(E)\right) + \frac{1}{2} \cdot \nu\left(g_1^{-1}(E)\right),$$

for all Borel sets $E \subset [0,1]$. We know that for

$$\pi_{\mathbf{t}} : \Sigma \to [0,1], \quad \pi_{\mathbf{t}}(\mathbf{i}) := \lim_{n\to\infty} g_{i_0} \circ \cdots \circ g_{i_n}(0),$$

we have

$$\nu = \nu_{\mathbf{t}} = (\pi_{\mathbf{t}})_* \mu,$$

where $\Sigma = \{0,1\}^{\mathbb{N}}$ and $\mu = \left\{\frac{1}{2}, \frac{1}{2}\right\}^{\mathbb{N}}$ Bernoulli measure on Σ.
Next, we can associate to the Furstenberg measure ν its entropy $h(\nu)$ and its Lyapunov exponent $\lambda(\nu)$, as before. By Birkhoff Ergodic Theorem we have

$$\lambda(\nu) := \int \log g_{i_0}'(\sigma \mathbf{i}) d\mu(\mathbf{i}).$$

The following follows as a special case of the Jaroszewska-Rams theorem, Theorem 3.

Theorem 24. *We have*

$$\dim_{\mathrm{H}}(\nu_{\mathbf{t}}) \leq \frac{h(\nu_{\mathbf{t}})}{-\lambda(\nu_{\mathbf{t}})}.$$

In particular, for those parameters $\mathbf{t}$ *for which* $h(\nu_{\mathbf{t}}) < -\lambda(\nu_{\mathbf{t}})$ *the measure* $\nu_{\mathbf{t}}$ *is singular.*

7.1. **The absolute continuity of the Furstenberg measure.** We want to find sufficient conditions for the Furstenberg measure to be absolutely continuous. We will find a parameter set U of positive Leb_4 measure such that for almost all parameters $\mathbf{t}$ from U the Furstenberg measure $\nu_{\mathbf{t}}$ is absolute continuous. To achieve this we apply the theory introduced by Simon, Solomyak and Urbanski c.f. [18] and [19] which says roughly speaking that if the (so called) transversality condition holds on a parameter domain then for Lebesgue almost parameters from this domain the measure $\nu_{\mathbf{t}}$ is absolute continuous whenever the ratio of entropy/Lyapunov exponent is greater than 1.
Throughout this section we use the notation

$$\text{for } i = 0, 1 \text{ we write } \widehat{i} := 1 - i.$$

and

$$\underline{d}_i := \min\left\{d_i, d_i^{-1}\right\}, \qquad \overline{d}_i := \max\left\{d_i, d_i^{-1}\right\}.$$

In the rest of the paper, the parameter $\mathbf{t}$ will be chosen from the open set $U \subset (\mathbb{R}^+)^4$ defined as follows:

Definition 25.

$$U := U_1 \cap U_2 \cap U_3,$$

where $U_1, U_2, U_3 \subset (\mathbb{R}^+)^4$ *are the set of parameters* $\mathbf{t} \in (\mathbb{R}^+)^4$ *for which the conditions in points* ***(1)***, ***(2)*** *and* ***(3)*** *below hold respectively.*

- **(1):** *The functions* g_0, g_1 *are monotone increasing strict contractions and*

$$g_0([0,1)) \cup g_1((0,1]) = [0,1] = [\mathrm{fix}(g_0), \mathrm{fix}(g_1)],$$

where $fix(g_0), fix(g_1)$ *are the unique fixed points of* g_0, g_1, *respectively. This is equivalent to*

$$0 < \alpha_i < \underline{d}_i \text{ for } i = 0, 1 \text{ and } \alpha_0 + \alpha_1 > 1$$

- **(2):** $g_{ii}([0,1]) \cap g_{\widehat{i}}([0,1]) = \emptyset$ *holds for i=0,1. Equivalently, both of the following two inequalities hold:*

$$(g_{00}(1) =) \frac{\alpha_0^2}{\alpha_0 + d_0(1-\alpha_0)} < 1 - \alpha_1 (= g_1(0))$$

and

$$(g_{11}(0) =)\frac{d_1(1-\alpha_1)+(1-\alpha_1)\alpha_1}{d_1(1-\alpha_1)+\alpha_1} > \alpha_0(= g_0(1)).$$

(3): *We write* $T^0(x) = x$ *is the identity, finally* $T^1(x) := 1 - x$. *We assume that either for* $i = 0$ *or for* $i = 1$:

$$\frac{\alpha_{\hat{i}} \cdot \overline{d}_{\hat{i}} \cdot \left(T^i \circ g_{\hat{i}}^{-1} \circ (T^i)^{-1}\right)(\alpha_i)}{\alpha_i} + \frac{\alpha_0 \cdot \alpha_1 \cdot \overline{d}_0 \cdot \overline{d}_1}{1 - \alpha_i \cdot \overline{d}_i} < 1 \tag{21}$$

holds. (This is a technical condition in order to check the so called transversality condition.)

To main result of this section describes properties of the measure in terms of its Lyapunov exponent.

Theorem 26. *For every* $d_0, d_1 \in \mathbb{R}^+$ *for* Leb_2 *almost all* $(\alpha_0, \alpha_1) \in [0,1]^2$ *if* $\mathbf{t} = (\alpha_1, \alpha_1, d_0, d_1) \in U$ *then*

(1) $\dim_{\mathrm{H}} \nu_{\mathbf{t}} = \min\left\{\frac{\log 2}{-\log \lambda}, 1\right\}$;

(2) *moreover, if* $\frac{\log 2}{-\log\lambda} > 1$ *then* $\nu_{\mathbf{t}}$ *is absolute continuous with respect to the Lebesgue measure.*

There are very few results about the absolute continuity of measures which are invariant with respect to a non-linear IFS. The only method available, up to our knowledge, is the theory developed in the papers [18] and [19]. In order to apply it we need to require that the maps are strict contractions (condition **(1)** of Definition 25) and that the transversality condition holds (mostly condition **(3)**). Even in the linear case (e.g. Bernoulli convolution measures) we can apply the transversality method to prove absolute continuity, only if the contractions are "small". This is corresponds to our condition **(2)**. In Definition 25 the first two conditions were stated first with a rather geometric description and then we expressed them in terms of formulas. The equivalence of these formulations are immediate. The only thing we need to observe to verify this is:

$$g_0'(x) = \frac{\alpha_0 d_0}{(x + d_0(1-x))^2} \text{ and } g_1'(x) = \frac{\alpha_1 d_1}{(d_1 x + (1-x))^2}$$

implies that

$$\alpha_i \underline{d}_i \le g_i'(x) \le \alpha_i \overline{d}_i \text{ for } i = 0, 1. \tag{22}$$

In the particular case that $d_0 = d_1 = 1$ then the maps g_0, g_1 are linear with slope α_0, α_1 respectively. Next we define the transversality condition and then we prove that conditions **(1)**, **(2)** and **(3)** of Definition 25 imply that the transversality condition holds.

Definition 27 (Transversality). *Let $d_0, d_1 \in \mathbb{R}^+$ be arbitrary fixed values. We say that the transversality condition holds on an open set of $V = V_{d_0,d_1} \subset [0,1]^2$ of parameters (α_0, α_1) if there exists $C > 0$ such that for every sequence $\mathbf{i} = (i_0, i_1, \dots), \mathbf{j} = (j_0, j_1, \dots) \in \Sigma$ with $i_0 \neq j_0$ we have:*

$$Leb_2\{(\alpha_0, \alpha_1) \in V : \mathbf{t} = (\alpha_0, \alpha_1, d_0, d_1),\ |\pi_{\mathbf{t}}(\mathbf{i}) - \pi_{\mathbf{t}}(\mathbf{j})| < r\} \leq Cr$$

for all $r > 0$.

It is easy to see (c.f. [18] and [19]) that in order to establish transversality we only need to do a computation involving the gradients:

Lemma 28. *Fix $d_0, d_1 \in \mathbb{R}^+$. Let $V \subset [0,1]^2$ be an open set. Assume that for all $(\alpha_0, \alpha_1) \in V$ for $\mathbf{t} = (\alpha_0, \alpha_1, d_0, d_1)$ and $\mathbf{i}, \mathbf{j} \in \Sigma$ with $\pi_{\mathbf{t}}(\mathbf{i}) = \pi_{\mathbf{t}}(\mathbf{j})$ and $i_0 \neq j_0$ we have*

$$||\nabla_{\alpha_0,\alpha_1}\pi_{\mathbf{t}}(\mathbf{i}) - \nabla_{\alpha_0,\alpha_1}\pi_{\mathbf{t}}(\mathbf{j})|| > 0, \tag{23}$$

where

$$\nabla_{\alpha_0,\alpha_1}\pi_{\mathbf{t}}(\mathbf{i}) := \left(\frac{\partial}{\partial\alpha_0}\pi_{\mathbf{t}}(\mathbf{i}), \frac{\partial}{\partial\alpha_1}\pi_{\mathbf{t}}(\mathbf{i})\right).$$

Then the transversality condition holds on V.

Now we can prove the main result of this section:

Proof of Theorem 26. Using [19, Theorem 7.2], to verify our theorem the only thing left to do is to prove that the conditions **(1)**, **(2)** and **(3)** imply that (23) holds. First we observe that the roles of the functions g_0 and g_1 are symmetric. Namely, for

$$\gamma_0(x) = \frac{\alpha_1 x}{x + d_1(1-x)} \text{ and } \gamma_1(x) = \frac{d_0 x + (1-\alpha_0)(1-x)}{d_0 x + (1-x)}.$$

we have

$$g_0(x) = T^1 \circ \gamma_1 \circ (T^1)^{-1}(x) \text{ and } g_1(x) = T^1 \circ \gamma_0 \circ (T^1)^{-1}(x). \tag{24}$$

That is, we obtain the graph of the function g_i by reflecting the graph of $\gamma_{\widehat{i}}$ to the center of the unit square. First we assume that condition **(3)** holds with $i = 0$. That is now we assume that

$$\frac{\alpha_1 \overline{d}_1 g_1^{-1}(\alpha_0)}{\alpha_0} + \frac{\alpha_0\alpha_1\overline{d}_0\overline{d}_1}{1-\alpha_0\overline{d}_0} < 1. \tag{25}$$

To see that this implies (23) first we verify that

$$\forall \mathbf{k} \in \Sigma \text{ we have } \left|\frac{\partial}{\partial\alpha_0}\pi_{\mathbf{t}}(\mathbf{k})\right| \leq \frac{1}{1-\alpha_0\overline{d}_0}. \tag{26}$$

Namely, we observe that for every $y \in [0,1]$ we have

$$\frac{\partial}{\partial \alpha_0} g_0(y) = \frac{1}{\alpha_0} \cdot g_0(y). \tag{27}$$

If $\frac{\partial}{\partial \alpha_0}\pi_{\mathbf{t}}(\mathbf{k}) \neq 0$ then there is an $n \geq 0$ such that the first zero in $\mathbf{k} = (k_0, k_1, \dots)$ is k_n. Then for $z = \pi_{\mathbf{t}}(\sigma^{n+1}\mathbf{k})$ we have

$$\begin{aligned} \frac{\partial}{\partial \alpha_0}\pi_{\mathbf{t}}(\mathbf{k}) &= \left(g_1^n\right)'(z) \cdot \left[\frac{1}{\alpha_0} g_0\left(g_0(z)\right) + g_0'(z) \cdot \frac{\partial}{\partial \alpha_0}\pi_{\mathbf{t}}(\sigma^{n+1}\mathbf{k})\right] \\ &\leq 1 + \alpha_0 \cdot \overline{d}_0 \cdot \frac{\partial}{\partial \alpha_0}\pi_{\mathbf{t}}(\sigma^{n+1}\mathbf{k}), \end{aligned}$$

where in the first line we used (27) and in the second line we used the facts that g_1 is a contraction, $\max_x g_0(x) = \alpha_0$ and (22). We obtain that (26) holds by induction. Next we fix arbitrary $\mathbf{i}, \mathbf{j} \in \Sigma$ with $\pi_{\mathbf{t}}(\mathbf{i}) = \pi_{\mathbf{t}}(\mathbf{j})$ and $i_0 \neq j_0$. Without loss of generality we may assume that $i_0 = 0$ $j_0 = 1$. Then by condition **(2)** we have $i_1 = 1$ and $j_1 = 0$. To verify (23) now we prove that

$$\frac{\partial}{\partial \alpha_0}(\pi_{\mathbf{t}}(\mathbf{i})) - \frac{\partial}{\partial \alpha_0}(\pi_{\mathbf{t}}(\mathbf{j})) > 0. \tag{28}$$

We define the real functions $h_1(\alpha_0) := \pi_{\mathbf{t}}(\sigma^2\mathbf{i})$ and $h_2(\alpha_0) := \pi_{\mathbf{t}}(\sigma^2\mathbf{j})$. We apply the chain rule, (22) and (27) twice to get

$$\frac{\partial}{\partial \alpha_0}(\pi_{\mathbf{t}}(\mathbf{i})) = \frac{\pi_{\mathbf{t}}(\mathbf{i})}{\alpha_0} + g_0'(g_1(h_1(\alpha_0))) \cdot g_1'(h_1(\alpha_0)) \cdot h_1'(\alpha_0) \geq \frac{\pi_{\mathbf{t}}(\mathbf{i})}{\alpha_0}$$

and

$$\begin{aligned} \frac{\partial}{\partial \alpha_0}(\pi_{\mathbf{t}}(\mathbf{j})) &= g_1'(g_0(h_2(\alpha_0))) \cdot \left(\frac{g_0(h_2(\alpha_0))}{\alpha_0} + g_0'(h_2(\alpha_0)) \cdot h_2'(\alpha_0)\right) \\ &\leq \alpha_1\overline{d}_1 \cdot \frac{g_1^{-1}(\pi_{\mathbf{t}}(\mathbf{i}))}{\alpha_0} + \frac{\alpha_0\overline{d}_0 \cdot \alpha_1\overline{d}_1}{1-\alpha_0\overline{d}_0}, \end{aligned}$$

where in the last line we used that $\pi_{\mathbf{t}}(\mathbf{i}) = \pi_{\mathbf{t}}(\mathbf{j})$. Thus

$$\frac{\partial}{\partial \alpha_0}\left(\pi_{\mathbf{t}}(\mathbf{i}) - \pi_{\mathbf{t}}(\mathbf{j})\right) > f(z),$$

where $f : [1-\alpha_1, \alpha_0] \to \mathbb{R}$ is defined as

$$f(z) := \frac{z}{\alpha_0} - \alpha_1\overline{d}_1 \cdot \frac{g_1^{-1}(z)}{\alpha_0} - \frac{\alpha_0\overline{d}_0 \cdot \alpha_1\overline{d}_1}{1-\alpha_0\overline{d}_0}.$$

By (22) the function f is strictly decreasing and by (25) we have that $f(z) \geq f(\alpha_0) > 0$ holds for all $1-\alpha_1 \leq z \leq \alpha_0$. Hence, (28) holds.

Now we turn to the case when (21) holds with $i = 1$. Using the first part of (24) this means that

$$\frac{\alpha_0 \overline{d}_0 \gamma_1^{-1}(\alpha_1)}{\alpha_1} + \frac{\alpha_0 \alpha_1 \overline{d}_0 \overline{d}_1}{1 - \alpha_1 \overline{d}_1} < 1. \tag{29}$$

Below we use that the Conditions **(1)** and **(2)** of Definition 25 remain unchanged if we interchange $(\alpha_0, \overline{d}_0)$ and $(\alpha_1, \overline{d}_1)$, so we can consider $\widetilde{\mathbf{t}} = (\alpha_1, \alpha_0, d_1, d_0)$ instead of $\mathbf{t} := (\alpha_0, \alpha_1, d_0, d_1)$. By definition the natural projection in this case will be

$$\begin{aligned} \pi_{\widetilde{\mathbf{t}}}(\mathbf{i}) &= \lim_{n\to\infty} \gamma_{i_1} \circ \gamma_{i_2} \circ \cdots \circ \gamma_{i_n}(0) \\ &= \lim_{n\to\infty} T^1(g_{\widehat{i}_1} \circ g_{\widehat{i}_2} \circ \cdots \circ g_{\widehat{i}_n}((T^1)^{-1}0)) = T^1\left(\pi_{\mathbf{t}}(\widehat{\mathbf{i}})\right). \end{aligned}$$

Note that assumption (29) for the IFS $\{\gamma_1, \gamma_2\}$ and $\widetilde{\mathbf{t}}$ is exactly the same as the assumption (25) for the IFS $\{g_0, g_1\}$ and $\mathbf{t}$. So, we can apply the previous argument to show that whenever $\pi_{\widetilde{\mathbf{t}}}(\mathbf{i}) = \pi_{\widetilde{\mathbf{t}}}(\mathbf{j})$ (which is equivalent to $\pi_{\mathbf{t}}(\widehat{\mathbf{i}}) = \pi_{\mathbf{t}}(\widehat{\mathbf{j}})$) with $i_1 = 0, j_1 = 1$ then

$$-\left(\frac{\partial}{\partial \alpha_1}(\pi_{\mathbf{t}}(\widehat{\mathbf{i}})) - \frac{\partial}{\partial \alpha_1}(\pi_{\mathbf{t}}(\widehat{\mathbf{j}}))\right) = \frac{\partial}{\partial \alpha_1}(\pi_{\widetilde{\mathbf{t}}}(\mathbf{i})) - \frac{\partial}{\partial \alpha_1}(\pi_{\widetilde{\mathbf{t}}}(\mathbf{j})) > 0.$$

This completes the proof of the Theorem. □

8. Examples of absolute continuity of the Furstenberg measure

We can restrict attention to considering matrices of the particular form

$$A_0 = \begin{pmatrix} \alpha_0 & 0 \\ 1-\alpha_0 & d_0 \end{pmatrix} \text{ and } A_1 = \begin{pmatrix} d_1 & 1-\alpha_1 \\ 0 & \alpha_1 \end{pmatrix}$$

where $0 < 1 - \alpha_1 < \alpha_0 < 1$ and $d_0, d_1 \in \mathbb{R}^+$, since, as we have seen, this can be achieved by an affine change of coordinates and since scaled matrices have the same projective actions. This gives a four dimensional parameter space.

If we assume for simplicity that $d_0 = d_1 = 1$ then the functions g_0, g_1 are linear: $g_0(x) = \alpha_0 x, g_1(x) = \alpha_1 x + 1 - \alpha_1$. The absolute continuity of the invariant measure for this linear case was investigated by Ngai, Wang [14]. (C.f. [14, Figure 1] for the domain of absolute continuity.) While in [14] only the linear case was considered, our calculations also work in the case of (d_0, d_1) which are sufficiently close to the linear case $d_0 = d_1 = 1$. The conditions **(1)** and **(2)** of Definition 25 can be summarized in the linear case $d_0 = d_1 = 1$ as follows:

$$\alpha_0^2 < 1 - \alpha_1 < \alpha_0 < 1 - \alpha_1^2.$$

Moreover, in the linear case, condition **(3)** of Definition 25 is as follows:

$$\text{either } \frac{\alpha_0+\alpha_1-1}{\alpha_0}+\frac{\alpha_0\alpha_1}{1-\alpha_0}<1 \text{ or } \frac{\alpha_1+\alpha_0-1}{\alpha_1}+\frac{\alpha_0\alpha_1}{1-\alpha_1}<1.$$

In particular, solutions include $\alpha_0 = 17/128$ and $\alpha_1 = 7/8$ and thus we could consider a full measure set of matrices in a sufficiently small (four or eight dimensional) neighborhood of

$$A_0 = \begin{pmatrix} 17/128 & 0 \\ 111/128 & 1 \end{pmatrix} \text{ and } A_1 = \begin{pmatrix} 1 & 1/8 \\ 0 & 7/8 \end{pmatrix}$$

We can also consider the set of parameters (α_0, α_1) satisfying the above inequalities and for a.e. (α_0, α_1) in this set we have that the Furstenberg measure is absolutely continuous.

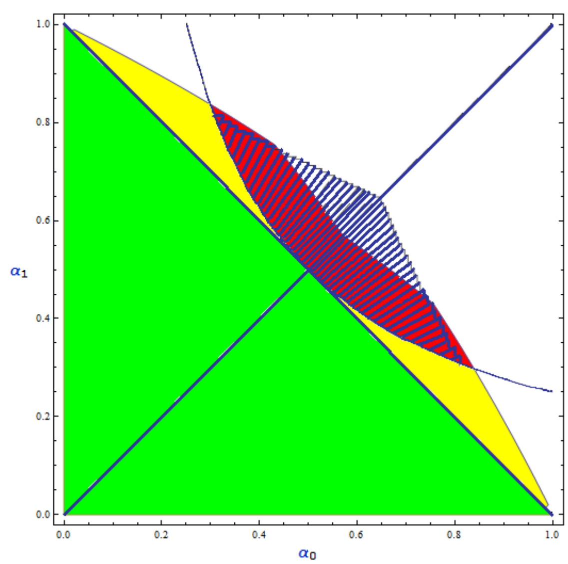

FIGURE 7. $d_0 = d_1 = 1$. In the green region dimension is equal to entropy over Lyapunov exponent. In the yellow region this remains true almost surely. In the red region we have absolute continuity. In the striped region the absolute continuity was proved by Ngai Wang.

For the Furstenberg measure, with probability vector $(p, 1-p)$, the entropy of the measure $\nu = (p, 1-p)^{\mathbb{Z}_+}$ is simply $h(\nu) = -p\log p - (1-p)\log(1-p)$. In particular, when $p = \frac{1}{2}$ this becomes $\log 2$. However, the Lyapunov exponent $\lambda(\nu)$ of the Furstenberg measure cannot be written in a closed form and can only be numerically estimated.

We recall that by definition $\lambda(\nu) = \lim_{n\to\infty} \lambda_n$, where

$$\lambda_n := \frac{1}{2^n} \sum_{i_0,\dots,i_{n-1}} \log \|A_{i_1} \cdots A_{i_n}\|.$$

The following table gives the values of these approximations. (It is also possible to get estimates on λ using cycle expansions.)

n	λ_n
1	0.287024757051608011
2	0.258808591315294742
3	0.245624281188549106
4	0.239666895599546522
5	0.237089539695933433
6	0.236022533561996868
7	0.235598268433514587
8	0.235435765643042602
9	0.235375877561900943
10	0.235354815592289776
11	0.235347897218792846
12	0.235345887479272859
13	0.235345459418261633
14	0.235345473898739984
15	0.235345571492518357

In particular, in this case we see that $\frac{h(\nu)}{\lambda(\nu)} = 2 \cdot 94523 \dots$. The value needed to be at least 1 to be consistent with the absolute continuity of (nearby) measures.

A non-linear example: We can also consider the case $d_0 = \frac{1}{d_1} = d \neq 1$. In this case g_0, g_1 are fractional linear functions. Unfortunately, the Lyapunov exponent of $\nu = (\frac{1}{2}, \frac{1}{2})^{\mathbb{Z}_+}$ is complicated to calculate. However, we can approximate it. By using the definition of Lyapunov exponent, we have

$$\lambda_\nu = -\int_\Sigma \log g'_{i_0}(\pi_{\mathbf{t}}(\sigma \mathbf{i})) d\nu(\mathbf{i}) =$$

$$-\log d - \frac{1}{2} \log \alpha_0\alpha_1 + 2 \int_\Sigma \log\left(\pi_{\mathbf{t}}(\mathbf{i}) + d(1 - \pi_{\mathbf{t}}(\mathbf{i}))\right) d\nu(\mathbf{i})$$

Since $\log(x + d(1-x))$ is monotone increasing or decreasing, depending on d, we can approximate the integral of it. For $d < 1$ by the following way

$$\sum_{|\mathbf{i}|=n} \frac{1}{2^n} \log\left(g_{\mathbf{i}}(0) + d(1 - g_{\mathbf{i}}(0))\right) \leq \int_{\Sigma} \log\left(\pi_{\mathbf{t}}(\mathbf{i}) + d(1 - \pi_{\mathbf{t}}(\mathbf{i}))\right) d\nu(\mathbf{i}) \leq \sum_{|\mathbf{i}|=n} \frac{1}{2^n} \log\left(g_{\mathbf{i}}(1) + d(1 - g_{\mathbf{i}}(1))\right)$$

We can see an approximation of the set of (α_0, α_1), where the Furstenberg measure is absolutely continuous for $d = 0.9$, on Figure 8.

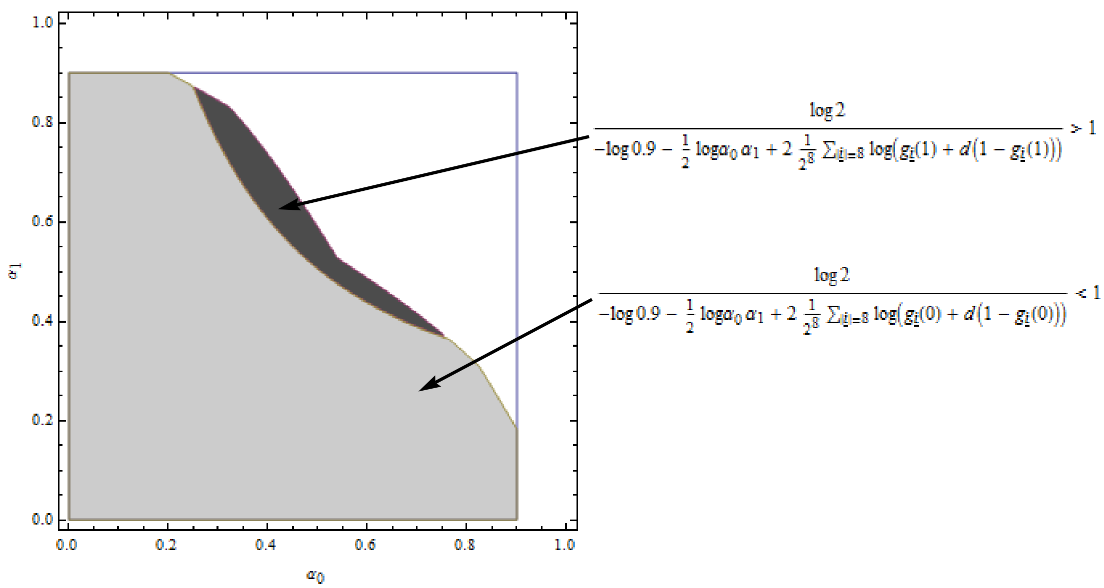

FIGURE 8. $d_0 = \frac{1}{d_1} = 0.9$. In the gray region, dimension is equal to entropy over Lyapunov exponent almost surely and the measure is singular. In the black region we have absolute continuity almost surely.

The absolute continuity domain include for example $\alpha_0 = 9/20$ and $\alpha_1 = 3/5$ and $d = d_0 = 1/d_1 = 9/10$ and thus we could consider a full measure set of matrices in a sufficiently small (four or eight dimensional) neighborhood of

$$A_0 = \begin{pmatrix} \frac{9}{20} & 0 \\ \frac{11}{20} & \frac{9}{10} \end{pmatrix} \text{ and } A_1 = \begin{pmatrix} \frac{10}{9} & \frac{2}{5} \\ 0 & \frac{3}{5} \end{pmatrix}.$$

where $\frac{h(\nu)}{\lambda(\nu)} > 1$. So the measure ν corresponding to the parameters which belong into this small neighborhood is almost surely absolute continuous.

9. Appendix

Here we briefly summarize the results of the papers [18] and [19] used in this note. We are given the family $\Phi^{\mathbf{t}} = \{\phi_1^{\mathbf{t}}, \dots, \phi_m^{\mathbf{t}}\}_{\mathbf{t}\in U}$ of hyperbolic IFS on the a compact interval $X \subset \mathbb{R}$. (Hyperbolic means that there exist $0 < c_1 < c_2 < 1$ such that for the $\mathcal{C}^2$ maps $\phi_k^{\mathbf{t}}$ we have $c_1 < |(\phi_k^{\mathbf{t}})'(x)| < c_2$ for all $k, \mathbf{t}, x$). We assume that the parameter domain $U \subset \mathbb{R}^d$ is a bounded open set with smooth boundary. As usual we denote the natural projection from $\Sigma := \{1, \dots, m\}^{\mathbb{N}}$ to X by

$$\pi_{\mathbf{t}}(\mathbf{i}) := \lim_{n\to\infty} \phi_{i_1}^{\mathbf{t}} \circ \cdots \circ \phi_{i_n}^{\mathbf{t}}(x),$$

where $x \in X$ is arbitrary. In this section we assume that we are given a σ-invariant ergodic measure μ on Σ and we always write

$$\nu_{\mathbf{t}} := \pi_{\mathbf{t}}\mu$$

for its push down measure. The next theorem shows how the dimension and absolute continuity of the measure $\nu_{\mathbf{t}}$ depends on the ratio of the entropy $h(\mu)$ of μ and the Lyapunov exponent $\lambda_{\mathbf{t}}$ of the measure $\nu_{\mathbf{t}}$ defined as

$$\lambda_{\mathbf{t}} := \int \phi_{i_1}^{\mathbf{t}}(\pi_{\mathbf{t}}(\sigma\mathbf{i}))d\mu(\mathbf{i}).$$

Below we often use the notation

$$\Sigma_2 := \{(\mathbf{i}, \mathbf{j}) \in \Sigma \times \Sigma : \ i_0 \neq j_0\}, \qquad f_{\mathbf{i},\mathbf{j}}(\mathbf{t}) := \pi_{\mathbf{t}}(\mathbf{i}) - \pi_{\mathbf{t}}(\mathbf{j}).$$

We recall for the reader how the following theorem is proved in [18] and [19]:

Theorem 29 (Simon, Solomyak and Urbanski)**.** *We assume that*

$$\text{(H)} \qquad \forall(\mathbf{i}, \mathbf{j}) \in \Sigma_2, \forall \mathbf{t} \in \overline{U} : \ \|\nabla f_{\mathbf{i},\mathbf{j}}(\mathbf{t})\| > 0.$$

Then

(1) *For* Leb_d *almost all* $\mathbf{t} \in U$, $\dim_{\mathrm{H}} \nu_{\mathbf{t}} = \min\left\{\frac{h(\mu)}{-\lambda_{\mathbf{t}}}, 1\right\}$*;*

(2) $\nu_{\mathbf{t}} \ll Leb$ *for* Leb_d *almost all* $\mathbf{t} \in \{\mathbf{t} \in U : \ -h(\mu)/\lambda_{\mathbf{t}} > 1\}$

We start with a Lemma from [18]. For a set $F \subset \mathbb{R}^d$ let $N_r(F)$ be the minimal number of balls needed to cover the set F.

Lemma 30. *(see* [18, Lemma 7.3]*) Let* $U \subset \mathbb{R}^d$ *be as above. Suppose that* f *is a* $\mathcal{C}^1$ *real-valued function defined in a neighborhood of the closure of* U *such that for some* $i \in \{1, \dots, d\}$ *there exists an* $\eta > 0$ *satisfying*

$$\text{(30)} \qquad \mathbf{t} \in U, \ |f(\mathbf{t})| \leq \eta \Longrightarrow \frac{\partial f(\mathbf{t})}{\partial t_i} \geq \eta.$$

Then there exists $C = C(\eta)$ *such that*

$$N_r(\{\mathbf{t} \in U : \ |f(\mathbf{t})| \le r\}) \le c \cdot r^{1-d}, \ \forall r > 0. \tag{31}$$

For every $(\mathbf{i}, \mathbf{j}) \in \Sigma_2$ we may apply this lemma for $f = f_{\mathbf{i},\mathbf{j}}(\mathbf{t})$ since (H) implies that (30) holds. Write $\eta_{\mathbf{i},\mathbf{j}}$ and $C_{\mathbf{i},\mathbf{j}}$ for the corresponding constants. For compactness $\underline{\eta} := \min_{(\mathbf{i},\mathbf{j}) \in \Sigma_2} \eta_{\mathbf{i},\mathbf{j}} > 0$ and $\overline{C} := \max_{(\mathbf{i},\mathbf{j}) \in \Sigma_2} C_{(\mathbf{i},\mathbf{j})} < \infty$. So,

$$\forall (\mathbf{i}, \mathbf{j}) \in \Sigma_2 : \qquad N_r(\{\mathbf{t} \in U : \ |f_{\mathbf{i},\mathbf{j}}| \le r\}) \le \overline{C} \cdot r^{1-d}$$

The last statement is called the strong transversality condition (c.f. [18, p. 454]) which clearly implies that the transversality condition holds: There exists a $\widetilde{c} > 0$ such that

$$\forall r > 0 : \ \mathrm{Leb}_d \{\mathbf{t} \in U : \ |f_{\mathbf{i},\mathbf{j}}| \le r\} \le \widetilde{c} \cdot r \tag{32}$$

Then we can apply [19, Theorem 7.2] which immediately yields the assertion of our theorem.

Acknowledgment. Research of Bárány was supported by the EU FP6 Research Training Network CODY and TAMOP - 4.2.2.B-10/1-2010-0009. Both Bárány and Simon were supported by grant OTKA # K71693.

References

[1] R. Bowen, *Equilibrium states and the ergodic theory of Anosov Diffeomorphisms* Springer-Verlag 1975.

[2] D. Blackwell, *The entropy of functions of finite-state Markov chains*, Trans. First Prague Conf. Information Theory, Statistical Decision Functions, Random Processes, (1957), 13-20.

[3] T. Cover, J. Thomas, *Elements if Information Theory*, Wiley, 1991 (second edition 2006).

[4] P. Cvitanovic, *Chaos - classical and quamtum* http://chaosbook.org/index.html

[5] , K. J. Falconer, *Techniques in Fractal Geometry*, Wiley, 1997.

[6] H. Furstenberg, *Non-commuting random products*, Transactions of the American Mathematical Society **108** (1963), 377-428.

[7] G. Han, B. Marcus, *Analyticity of entropy rate of hidden Markov chains*, IEEE Trans. Inform. Theory **52** (2006), no. 12, 5251–5266.

[8] G. Han, B. Marcus, *Derivates of Entropy Rate in Special Families of Hidden Markov Chains*, IEEE Transactions on Information Theory **53** (2007), 2642-2652.

[9] J. E. Hutchinson, *Fractals and self-similarity*, Indiana University Mathematics Journal **30**, (1981), 713-747.

[10] J. Jaroszewska, M. Rams, *On the Hausdorff dimension of invariant measure of weakly contracting on average measurable Iterated Function Scheme*, Journal of Stat. Phys. to appear.

[11] V. Kaimanovich, V. Le Prince, *Matrix random products with singular harmonic measure*, Preprint http://arxiv.org/abs/0807.1015.

[12] J. Marklof, Y. Tourigny and L. Wolowski, *Explicit invariant measures for products of random matrices*, Transactions of the AMS **360** (2008), 3391-3427.

[13] R.D. Mauldin, M. Urbański *Graph directed Markov Systems: geometry and dynamics of limit sets* Cambridge 2003.

[14] Sz.M. Ngai, Y. Wang, *Self-similar measures associated with ifs with non-uniform contraction ratios* Asian J. Math. **9** (2005), 227-244.

[15] E. Ordentlich, T. Weissman, *On the Optimality of Symbol by Symbol Filtering and Deoising* IEEE Transactions on Information Theory **52** (2006), 19-40.

[16] Y. Peres, *Domain of analytic continuation for the top Lyapunov exponent*, Ann. Inst. H. Poincar Probab. Statist. **28** no.1 (1992), 131–148.

[17] E. Seneta, *Non-negative matrices and Markov chains*, Springer, 1981.

[18] K. Simon, Solomyak and M. Urbanski, *Hausdorff dimension of limit sets for parabolic IFS with overlaps*, Pacific J. Math. 201 (2001), no. 2, 441–478

[19] K. Simon, Solomyak and M. Urbanski, *Invariant measures for parabolic IFS with overlaps and random continued fractions*, Trans. Amer. Math. Soc. 353 (2001), no. 12, 5145–5164

[20] `http://reference.wolfram.com/mathematica/tutorial/ConstrainedOptimizationGlobalNumerical.html`