PROJECTIONS OF THE RANDOM MENGER SPONGE

KÁROLY SIMON\textsuperscript{1,2,3} AND VILMA ORGOVÁNYI\textsuperscript{1,3}

Abstract. Using a similar random process to the one which yields the fractal percolation sets, starting from the deterministic Menger sponge we get the random Menger sponge. We examine its orthogonal projections from the point of Hausdorff dimension, Lebesgue measure and existence of interior points.

We obtain these results as special cases of our theorems stated for random self-similar IFSs. These are obtained by a random process similar to the fractal percolation, applied for the cylinder sets of a deterministic self-similar IFS, as in [9]. In this paper the associated deterministic IFS on the line is of the special form $S = \{ \frac{1}{L}x + t_i \}_{i=1}^m$, where $L \in \mathbb{N}$, $L \geq 2$ and $t_i \in \mathbb{Q}$.

Contents

1. Introduction \hspace{1cm} 2
1.1. Brief summary \hspace{1cm} 2
1.2. Pictural explanation of our results about $\mathcal{M}_p$ \hspace{1cm} 4
1.3. Notations \hspace{1cm} 5
1.4. Results \hspace{1cm} 9
2. Menger sponge \hspace{1cm} 12
2.1. Proof of Theorem 1.14 \hspace{1cm} 13
2.2. The dual nature of $\mathcal{M}_p$ \hspace{1cm} 14
2.3. Proof of Theorem 1.15 assuming Proposition 2.4 \hspace{1cm} 15
2.4. Proof of Proposition 2.4 \hspace{1cm} 16

\textsuperscript{1}Department of Stochastics, Institute of Mathematics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary

\textsuperscript{2}Alfréd Rényi Institute of Mathematics – Eötvös Loránd Research Network, Reáltanoda u. 13-15., 1053 Budapest, Hungary

\textsuperscript{3}MTA-BME Stochastics Research Group, Műegyetem rkp. 3., H-1111 Budapest, Hungary

E-mail addresses: simonk@math.bme.hu, orgovanyi.vilma@gmail.com.

Date: 9th May 2022.
1. INTRODUCTION

1.1. Brief summary. The Mandelbrot percolation Cantor set is a two-parameter family of random sets on \( \mathbb{R}^d \). Namely, fix the parameters \( K \geq 2 \) and \( p \in (0, 1) \). In the first step of the construction we partition the unit cube \([0, 1]^d\) into axes-parallel cubes of side length \( 1/K \). Each of these cubes are retained with probability \( p \) and discarded with probability \( 1 - p \) independently. This step is repeated independently in each of the retained cubes ad infinitum or until no retained cubes left. The random set we end up with is the Mandelbrot percolation set.

Falconer and Jin introduced a generalization of the Mandelbrot percolation Cantor sets [9]. In this paper we consider a special case of Falconer and Jin’s construction.

Namely, we apply the random process, which results in the Mandelbrot percolation set, on the subsequent level cylinders in the construction of a self-similar set.

More precisely, to understand the nature of the above-mentioned random process, we consider a (deterministic) \( M \)-array tree \( \mathcal{T} \). That is every node of \( \mathcal{T} \) has exactly \( M \) children. (In the construction of the Mandelbrot percolation set \( M = K^d \).) We assign a random label (from \( \{0, 1\} \)) to each of these nodes. The label of the root \( \emptyset \) is equal to 1 and the random label of all other nodes are independent Bernoulli(\( p \)) random variables. A level \( n \) node is retained if all of its ancestors are labelled with 1. In the mandelbrot percolation example, every retained level-\( n \) node naturally corresponds to a retained level \( n \) cube. An infinite path starting from the root is retained if all the nodes of the path are labelled with 1. It may happen that no infinite paths are retained. This
event is called extinction. The set of retained level \(n\) nodes is denoted by \(\mathcal{E}_n\) for an \(n \in \mathbb{N} \cup \{\infty\}\). In the case of the Mandelbrot percolation, every element of \(\mathcal{E}_\infty\) naturally correspond to a point of the Mandelbrot percolation Cantor set.

A generalization of the Mandelbrot percolation sets can be obtained if we consider a self-similar IFS (Iterated Function System) \(\mathcal{F} := \{f_i\}_{i=1}^M\) on \(\mathbb{R}^d\) and, we retain the elements of the attractor of \(\mathcal{F}\) having a symbolic representation from \(\mathcal{E}_\infty\). We call these random sets coin tossing self-similar sets since we decide if a cylinder set is retained or not as a result of subsequent coin-tossing. For more detailed notations see Definition 1.1.

If the self-similar IFS is the one whose attractor is the so-called Menger sponge then this random construction yields the random Menger sponge. More precisely, first we define the (deterministic) Menger sponge, which is the natural three-dimensional generalization of the well-known Sierpiński-carpet. We partition the unit cube \([0,1]^3\) into 27 cubes of side length \(1/3\).

For each of the 6 faces of \([0,1]^3\) there is a central cube. We remove these 6 cubes as well as the cube which contains the center of \([0,1]^3\). We retain the remaining 20 cubes. Then we repeat this process in each of the retained cubes ad infinitum. The set which remains after infinitely many steps is the Menger sponge \(\mathcal{M}\). See Figure 1 for the first four approximations of \(\mathcal{M}\).

To construct a random Menger sponge we need a parameter \(p \in (0,1)\). Instead of retaining all of the 20 cubes we retain each with probability \(p\) and discard it with probability \(1 - p\) independently. Then we repeat this random construction in each retained cube ad infinitum or until there is no retained cubes left since it is possible that in finitely many steps we have no cubes retained. It follows from the theory of Branching Processes that whenever \(p > \frac{1}{20}\) we end up with a non-empty Cantor set with positive probability. This
random set is called random Menger sponge corresponding to probability $p$ and it is denoted by $\mathcal{M}_p$. For the precise definition see Example 1.4. We study the structure of projections of $\mathcal{M}_p$ to lines trough the origin of the form $\text{proj}_{(a,b,c)}(x, y, z) := ax + by + cz$. We say that $\text{proj}_{(a,b,c)}$ is rational if $a, b, c \in \mathbb{Q}$. It follows from Mattila’s Theorem [11] that for Lebesgue almost all $(a, b, c)$ the projection of $\mathcal{M}_p$

1. has Hausdorff dimension $\min\{1, \dim_H(\mathcal{M}_p)\}$,
2. if $\dim_H(\mathcal{M}_p) > 1$ then $\mathcal{L}(\text{proj}_{(a,b,c)}\mathcal{M}_p) > 0$.

Here we prove results of similar nature but ones which are much more precise in this special case.

1.2. Pictural explanation of our results about $\mathcal{M}_p$. Our results regarding the random Menger sponge (Theorems 1.14-1.19) are summarized on Figure 2. Many of these are special cases of theorems that we prove in this paper for more general coin tossing self-similar IFSs. A number of results presented on Figure 2, are about the $(1,1,1)$-projection of $\mathcal{M}_p$. We remark that corresponding results can be proved for rational projections satisfying some additional conditions. These conditions can be easily checked for the $(1,1,1)$-projection and in this case, we can determine the parameter intervals where these additional conditions hold.

1.2.1. Explanation of Figure 2. Under the baseline there is information about the proj$_{(1,1,1)}$-projected set. Information about the random Menger sponge itself are visualized by colored rectangles arround the $p$-axis, and lastly the information about the proj$_{(1,0,0)}$-projection (i.e. the projection to the coordinate axis) are placed above the baseline.

In what follows we present some further explanation to this figure. First consider the proj$_{(1,1,1)}$-projection of $\mathcal{M}_p$. It follows from Theorem 1.19 that there exists a $B \in (0.15, 0.1514\ldots)$ such that:

a) For $p \in (0.15, B)$ the Lebesgue-measure of proj$_{(1,1,1)}(\mathcal{M}_p)$ is zero almost surely, by Theorem 1.16, despite the fact that $\mathcal{M}_p$ has Hausdorff dimension greater than 1 almost surely conditioned on non-extinction by Fact 1.13.

b) For $p \in (B, 0.1666\ldots)$, conditioned on non-extinction, proj$_{(1,1,1)}(\mathcal{M}_p)$ is a set of positive Lebesgue measure (Theorem 1.19) which contains no interior points (Theorem 1.18) almost surely.
For $p \in (0.166\ldots, 1]$, conditioned on non-extinction, $\text{proj}_{(1,1,1)}(\mathcal{M}_p)$ contains an interval almost surely by Theorem 1.18.

About the other projections of the random Menger sponge $\mathcal{M}_p$ we can say the following:

d) Let $p \in (0.25, 1]$. Then simultaneously for all directions $(a, b, c)$ the projection contains an interval almost surely conditioned on non-extinction by Theorem 1.15.

This is interesting when $p < 0.35\ldots$ since in this case the Menger sponge itself is totally disconnected by Theorem 1.14. Also the bound 0.25 is sharp because for $p < 0.25$ the $\text{proj}_{(1,0,0)}$-projection does not contain an interval almost surely, by Theorem 1.17. As we already mentioned the results above are special cases of our more general theorems, which are presented in Section 1.4. To state them we need to introduce our basic notations.

1.3. Notations. Before we give the precise definition of the coin tossing self-similar sets, first we define the deterministic self-similar sets in $\mathbb{R}^d$. Given a self-similar IFS $\mathcal{F}$ on $\mathbb{R}^d$

\begin{equation}
\mathcal{F} := \{ f_i(x) := r_i Q_i x + t_i \}_{i=0}^{M-1}, \quad f_i : \mathbb{R}^d \to \mathbb{R}^d, \quad r_i \in (0,1), \quad Q_i \in O(d), \quad t_i \in \mathbb{R}^d.
\end{equation}

We use the short hand notations

\begin{equation}
f_{i_1,\ldots,i_n} := f_{i_1} \circ \cdots \circ f_{i_n}, \quad r_{i_1,\ldots,i_n} := r_{i_1} \cdots r_{i_n}, \quad [M] := \{0,\ldots,M-1\}.
\end{equation}
It is easy to see that we can choose

\[(1.3) \quad B \subset \mathbb{R}^d, \text{ compact such that } f_i(B) \subset B \text{ for all } i \in [M].\]

Then the union of all \(n\)-cylinders \(\bigcup_{i_1, \ldots, i_n} f_{i_1, \ldots, i_n}(B)\) form a nested sequence of compact sets. Their intersection is the attractor

\[\Lambda := \bigcap_{n=1}^{\infty} \bigcup_{i_1, \ldots, i_n} f_{i_1, \ldots, i_n}(B).\]

The definition of \(\Lambda\) does not depend on the choice of \(B\) as long as \(B\) satisfies \(1.3\).

**Definition 1.1** (Coin tossing self-similar sets). Let \(F := \{f_i\}_{i=0}^{M-1}\) be a (deterministic) self-similar IFS on \(\mathbb{R}^d\) as it was defined in \((1.1)\) and let \(p \in (0, 1)\).

The corresponding coin tossing self-similar set \(\Lambda_F(p)\) is defined as follows: In the first step for every \(k \in [M]\) we toss (independently) a bias coin which lands on head with probability \(p\). The random subset \(X_1 \subset [M]\) consists of those \(k \in [M]\) for which the coin tossing resulted in head. Assume that we have already constructed \(X_n \subset [M]^n\). Then for every node \(i \in X_n\) we define (independently of everything) the random set \(X_i^1 \subset [M]\) which has the same distribution as \(X_1\). The set of the offsprings of \(i\) is defined by \(O(i) = \{ik \in [M]^{n+1} : k \in X_i^1\}\), where \(ik = i_1, \ldots, i_n, k\) if \(i = i_1, \ldots, i_n\). Finally, we form \(X_{n+1} = \bigcup_{i \in X_n} O(i) \subset [M]^{n+1}\). Then the coin tossing self-similar set is defined by

\[(1.4) \quad \Lambda_F(p) := \bigcap_{n=1}^{\infty} \bigcup_{i \in X_n} f_i(B),\]

where \(B\) is chosen as in \((1.3)\).

We do not assume that the Open Set Condition (OSC) (see [7]) holds for \(F\). However, we mention the following theorem.

**Theorem 1.2** (Falconer [8], Mauldin-Williams [12]). Let \(F\) be a deterministic self-similar IFS, as in Definition 1.1, which satisfies the OSC. Then almost surely, conditioned on non-extinction,

\[(1.5) \quad \dim_H \Lambda_F(p) = \dim_B \Lambda_F(p) = s, \quad \text{where } \sum_{i=0}^{M-1} pr_i^s = 1,\]

where \(r_i\) is the contraction ratio of the similarity mapping \(f_i\).
Motivated by this formula we introduce the similarity dimension of a coin tossing self-similar set $\Lambda_F(p)$:

$$\dim_{\text{Sim}} \Lambda_F(p) := s, \text{ where } \sum_{i=0}^{M-1} p r_i^s = 1.$$  

In this paper we only consider homogeneous coin tossing self-similar IFSs, which means that all contraction ratios $r_i$ are equal to the same $r \in (0, 1)$. In this homogeneous case, formula (1.5) simplifies to

$$(1.7) \quad (\text{OSC & } r_i \equiv r, \forall i) \implies s = \dim_{\text{H}} \Lambda_F(p) = \dim_{\text{B}} \Lambda_F(p) = \frac{\log(Mp)}{-\log r},$$

almost surely, conditioned on non-extinction.

A more detailed definition of a coin tossing self-similar set, which describes the ambient probability space can be found in Section 3.1.

**Example 1.3** ((Homogeneous) Mandelbrot percolation). The homogeneous Mandelbrot percolation on $\mathbb{R}^d$ with parameters $(K, p)$ can be obtained as a special case of the construction defined above by choosing $M = K^d$ and

$$\mathcal{F} := \left\{ f_i(x) = \frac{1}{K} x + t_i \right\}_{i=0}^{M-1},$$

where $\{t_i\}_{i=0}^{M-1}$ is an enumeration of the left bottom corners of the $K$-mesh cubes contained in $[0, 1]^d$.

**Example 1.4** (Random Menger sponge). The (deterministic) Menger sponge is the attractor (see Figure 1) of the following self-similar IFS in $\mathbb{R}^3$:

$$\mathcal{F} := \left\{ f_i(x) = \frac{1}{3}(x) + t_i \right\}_{i=0}^{19},$$

where $\{t_i\}_{i=0}^{19}$ is an enumeration of the following set

$$\left\{ 0, \frac{1}{3}, \frac{2}{3} \right\} \times \left\{ \left( \frac{1}{3}, \frac{1}{3}, 0 \right), \left( \frac{1}{3}, 0, \frac{1}{3} \right), \left( 0, \frac{1}{3}, \frac{1}{3} \right), \left( \frac{1}{3}, \frac{1}{3}, \frac{2}{3} \right), \left( \frac{1}{3}, \frac{2}{3}, \frac{1}{3} \right), \left( \frac{2}{3}, \frac{1}{3}, \frac{1}{3} \right), \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right) \right\}.$$

We obtain the random Menger sponge by applying the random construction introduced in Definition 1.1 for the deterministic IFS above. As we have already mentioned we denote the random Menger sponge with parameter $p$ with $\mathcal{M}_p$.

If we project the random Menger sponge to straight lines, then the resulting random set is a coin tossing self-similar set on the line.
**Example 1.5** (Coin tossing integer self-similar sets on the line). We obtain the coin tossing integer self-similar sets on the line by applying the random construction introduced in Definition 1.1 for the following deterministic IFS:

\[
\mathcal{F} := \left\{ f_i(x) := \frac{1}{L} x + t_i \right\}_{i=0}^{M-1}, \quad f_i : \mathbb{R} \to \mathbb{R}, \quad L \in \mathbb{N} \setminus \{0, 1\}, \quad t_i \in \mathcal{N},
\]

where \( \mathcal{N} \subset \mathbb{R} \) is a lattice.

**Remark 1.6.** Without loss of generality (see [4, Section 1.3.3]) we may assume, that

\[
\mathcal{N} := \mathbb{N}, \quad 0 = t_0 \leq t_1 \leq t_2 \leq \cdots \leq t_{M-1}, \quad \text{and} \quad t_{M-1}|L - 1.
\]

**Remark 1.7.** In the deterministic case usually we require that the elements of the IFS \( \mathcal{F} \) are different. However, in the random case we allow repetition among the functions of \( \mathcal{F} \). This is reasonable since even if \( f_i = f_j \) we randomize them differently. For example, if \( \mathcal{F} := \left\{ \frac{1}{3}x, \frac{1}{3}x, \frac{1}{3}x + \frac{2}{3} \right\} \) and \( \mathcal{S} := \left\{ \frac{1}{3}x, \frac{1}{3}x + \frac{2}{3} \right\} \), then the coin tossing self-similar sets \( \Lambda_{\mathcal{F}}(p), \Lambda_{\mathcal{S}}(p) \) are different.

**Definition 1.8.** Given the deterministic IFS \( \mathcal{F} := \{ f_i \}_{i=0}^{M-1} \) we select all distinct elements of \( \mathcal{F} \) and form a new IFS from these functions, we denote it by \( \mathcal{S} := \{ S_i \}_{i=0}^{m-1} \subset \mathcal{F} \). That is for every \( i \in [M] \) there exists a unique \( j \in [m] \) such that \( f_i = S_j \). For every \( j \in [m] \) let

\[
n_j := \# \{ f_i \in \mathcal{F} : f_i = S_j \}
\]

and we define the probability vector

\[
q := (q_1, \ldots, q_m), \quad \text{where} \quad q_j := \frac{n_j}{M}.
\]

For a \( k \geq 2 \) we write \( \Sigma^{(k)} := [k]^N \), where we remind the reader that we defined \( [k] := \{0, 1, \ldots, k - 1\} \).

We introduce the natural (probability) measure \( \mu := q^N \) on \( \Sigma^{(m)} \). We define the natural projection \( \Pi^{(m)} : \Sigma^{(m)} \to \mathbb{R} \)

\[
\Pi^{(m)}(i) := \lim_{n \to \infty} S_{i_1, \ldots, i_n}(0),
\]

for \( i = i_1, \ldots, i_n \in \Sigma^{(m)} \). The push forward of the natural measure \( \mu \) is denoted by \( \nu \),

\[
\nu := \Pi^{(m)}_* \mu.
\]
Following Ruiz [15, Section 3.1.1] we define the L-adic intervals:

\[ D_k := \left\{ \left[ (i-1) \cdot L^{-k}, i \cdot L^{-k} \right] : i \in \mathbb{Z} \right\}, \quad \text{for } k \in \{-1, 0, \ldots \}. \]

Since \( \nu \) is compactly supported there exist finitely many intervals, called basic types,

\[ J^0, \ldots, J^{N-1} \in D_{-1} \] such that \( \text{spt}(\nu) \subset \bigcup_{i \in [N]} J^i \) and

\[ \nu(J^i) > 0 \text{ for every } i \in [N], \]

here we assume that the intervals \( J^0, \ldots, J^{N-1} \) are arranged in an increasing order. It is clear from definition that the interval spanned by the attractor is \( I := \left[ 0, L^{\frac{t_m-1}{L-1}} \right] \). It follows from our assumption \( L - 1 \mid t_{m-1} \) that the right endpoint of \( J^{N-1} \) coincides with the right endpoint of \( I \). For every \( k \in [N] \) the interval \( J^k \) subdivides into \( L^n \) intervals from \( D_{n-1} \) (of length \( L^{-(n-1)} \)) which are denoted by \( J^k_a = J^k_{a_1, \ldots, a_n} \) for \( a = (a_1, \ldots, a_n) \in [L]^n \). We define the \( N \times N \) matrices \( \{ A_a \}_{a=0}^{L-1} \):

\[ A_a(\ell, k) := \begin{cases} n_i, & \text{if } \exists i \quad S_i(J^k) = J^\ell_a; \\ 0, & \text{otherwise.} \end{cases} \]

Note that we index the rows and columns of the matrices \( A_a \) from 0 to \( N-1 \). For \( a = (a_1, \ldots, a_n) \in [L]^n \),

\[ A_a := A_{a_1} \cdots A_{a_n}. \]

The \( j \)-th column sum (CS) of the \( a \)-th matrix is denoted by

\[ CS_{a,j} := \sum_{i=0}^{N-1} A_a(i, j). \]

For the rest of this section let \( \Lambda_{\mathcal{F}}(p) \) denote the coin tossing integer self-similar set on the line, defined in Example 1.5 with parameters \( \mathcal{F} \) and \( p \).

### 1.4. Results

In the rest of this Section we deal with coin tossing integer self-similar IFSs \( \mathcal{F} \) on the line introduced in Example 1.5 including Remark 1.6. That is for the rest of this Section we always assume that

\[ \mathcal{F} := \left\{ f_i(x) := \frac{1}{L} x + t_i \right\}_{i=0}^{M-1}, \]

\[ L \in \mathbb{N} \setminus \{0, 1\}, \quad t_i \in \mathbb{N}, 0 = t_0 \leq \cdots \leq t_{M-1}, \quad L - 1 \mid t_{M-1}. \]

Moreover, throughout this Section we use the notation introduced in Section 1.3. The new results of this paper are as follows:
Theorem 1.9. The coin tossing integer self-similar set $\Lambda_{\mathcal{F}}(p)$ contains an interval, almost surely conditioned on non-extinction, if the following two conditions hold

1. $p \cdot CS_{a,U} > 1$ for all $a \in [L]$ and $U \in [N]$. That is $p$ is larger than the reciprocal of every column sum of every matrix.
2. There exists a $b \in [L]^*$ and $U \in [N]$ such that $A_b(U,V) > 0$ for all $V \in [N]$. That is there exist a product $(A_b, b \in [L]^*)$ of the matrices with a strictly positive row.

Theorem 1.10. The coin tossing integer self-similar set $\Lambda_{\mathcal{F}}(p)$ does not contain any intervals, almost surely, if there exists an $a \in [L]$ such that the spectral radius of the matrix $p \cdot A_a$ is smaller than 1.

Theorem 1.11. The coin tossing integer self-similar set $\Lambda_{\mathcal{F}}(p)$ has positive Lebesgue measure, almost surely conditioned on non-extinction, if the following two conditions hold:

1. $p \cdot L \cdot (\prod_{a=0}^{L-1} CS_{a,U}) > 1$ for all $U \in [N]$. That is for every $U \in [N]$ column index we consider the geometric mean of the $U$-th column sums of the matrices $\{A_a\}_{a=0}^{L-1}$ and we denote it by $g_U$. Our assumption is that $p > \max_{U \in [N]} \frac{1}{g_U}$.
2. For every $b \in [L]$ there exists an $U \in [N]$ such that $A_b(U,V) > 0$ for all $V \in [N]$. That is, every matrix $A_b (b \in [L])$ has a positive row.

Theorem 1.12. There exists a $B > \frac{L}{M}$ such that for $p \in (0, B)$ the Lebesgue measure of the coin tossing integer self-similar set $\Lambda_{\mathcal{F}}(p)$ is almost surely 0.

The relevance of the bound $\frac{L}{M}$ is that the similarity dimension $\dim_{\text{Sim}} \Lambda_{\mathcal{F}}(p) > 1$ if and only if $p > \frac{L}{M}$. Thus, for $p \in \left(\frac{L}{M}, B\right)$ the similarity dimension of $\Lambda_{\mathcal{F}}(p)$ is greater than 1 but its Lebesgue measure is zero almost surely.

In the following we state theorems about the random Menger sponge (defined in Example 1.4) and its rational projections (see (1.16)). As we mentioned earlier, part of them are consequences of the three theorems above, but some are special to the random Menger sponge $\mathcal{M}_p$. It is clear that the IFS which defines the Menger sponge (see Example 1.4) satisfies the OSC. So, we can apply formula (1.7), with the substitution $r = 1/3$ and $M = 20$. This yields that

Fact 1.13. For almost all realizations conditioned on non-extinction we have $\dim_H \mathcal{M}_p = \frac{\log 20 \cdot p}{\log 3}$, which is greater than 1 if and only if $p > \frac{3}{20} = 0.15$. 

PROJECTIONS OF THE RANDOM MENDER SPONGE 10
Now we state the theorems regarding the random Menger sponge. The following theorems are visualized on Figure 2. The contents of these theorems have already been mentioned in Section 1.2. Here we just give a more precise formulation.

Recall, that $\text{proj}_{(a,b,c)}(x, y, z) : \mathbb{R}^3 \to \mathbb{R}$ denote the scalar multiple of the orthogonal projection to the vector $(a, b, c)$,

$$
\text{proj}_{(a,b,c)}(x, y, z) := ax + by + cz,
$$

and the projection is rational if $a, b, c \in \mathbb{Q}$.

**Theorem 1.14.** If $p < \frac{1}{\sqrt{8}} = 0.35 \ldots$, then $\mathcal{M}_p$ is totally disconnected almost surely conditioned on non-extinction.

**Theorem 1.15.** For any projection $(\text{proj}_{(a,b,c)})$ if $p > 0.25$ then $\text{proj}_{(a,b,c)}(\mathcal{M}_p)$ contains an interval almost surely conditioned on non-extinction.

**Theorem 1.16.** For every rational projection $\text{proj}_{(a,b,c)}$ there exists a $p' = p'(a, b, c) > \frac{3}{20}$ such that for every $p < p'$: $\text{Leb}_1(\text{proj}_{(a,b,c)}(\mathcal{M}_p)) = 0$ almost surely.

**Theorem 1.17.** If $p < 0.25$ then the projection of the random Menger sponge onto the $x$-axis $(\text{proj}_{(1,0,0)}(\mathcal{M}_p))$ does not contains an interval almost surely.

**Theorem 1.18.** If $p > \frac{1}{6} = 0.1666 \ldots$, then $\text{proj}_{(1,1,1)}(\mathcal{M}_p)$ contains an interval almost surely conditioned on non-extinction, and if $p < \frac{1}{6}$, then $\text{proj}_{(a,b,c)}(\mathcal{M}_p)$ does not contain an interval almost surely.

**Theorem 1.19.** If $p > (8 \cdot 6)^{-\frac{1}{3}} = 0.1514 \ldots$, then $\text{Leb}_1(\text{proj}_{(1,1,1)}(\mathcal{M}_p)) > 0$ almost surely conditioned on non-extinction.

1.4.1. *The organization of the rest of the paper.* Firstly in Section 2 we prove the theorems regarding the random Menger sponge (Theorem 1.14-1.19). Theorem 1.16-1.19 are special cases of Theorem 1.9-1.12, hence we prove them assuming those more general theorems. For Theorem 1.14 we use a standard branching process type argument. To prove Theorem 1.15 we use the result of Simon and Vágó [16].

In section 3 we present the proof of the general theorems, Theorem 1.9-1.12. We start with proving Theorem 1.9 using a similar argument as in [6], secondly we prove Theorem 1.10 by a standard argument, then we prove Theorem 1.11 in a similar way as in [13]. Finally, we prove Theorem 1.12 by firstly considering the deterministic attractor in a similar way as in [3]
in a similar setup as in [15]. Using this deterministic result we can prove Theorem 1.11.

In the Appendix we present another example, namely two of the projections of random Sierpiński carpet. We briefly summarize the related results, and add some new by applying our theorems. We then compare the results regarding the random Menger sponge and the random Sierpiński carpet.

2. Menger sponge

We consider the projections \( \text{proj}_{(1,0,0)} \) and \( \text{proj}_{(1,1,1)} \) of \( \mathcal{M}_p \). Using that all rational projections of \( \mathcal{M}_p \) (after proper re-scaling) can be considered as coin tossing integer self-similar sets, we can use the general theorems (Theorem 1.9-1.12) for the proofs of Theorems 1.16-1.19.

Example 2.1 (The \( \text{proj}_{(1,0,0)} \) projection of the random Menger sponge). We project the random Menger sponge to the x-axis, hence the corresponding IFS is the following (see Definition 1.8): \( S_i(x) = \frac{1}{3}x + t_i, \ t_i \in \{0, 3, 6\} \), with \( (n_0, n_1, n_2) = (8, 4, 8) \), and basic types (see (1.17)) \( J^0 = [0, 3], \ J^1 = [3, 6], \ J^2 = [6, 9] \). Hence, the corresponding matrices (see (1.18)) are,

\[
B_0 = \begin{bmatrix} 8 & 0 & 0 \\ 4 & 0 & 0 \\ 8 & 0 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 & 8 & 0 \\ 0 & 4 & 0 \\ 0 & 8 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 & 0 & 8 \\ 0 & 0 & 4 \\ 0 & 0 & 8 \end{bmatrix}.
\]

Figure 3. Illustration of Example 2.2

Example 2.2 (The \( \text{proj}_{(1,1,1)} \) projection of \( \mathcal{M}_p \)). \( S_i(x) = \frac{1}{3}x + i, \ i \in [7] \). \( (n_0, \ldots, n_6) = (1, 3, 3, 6, 3, 3, 1) \). \( J^0 = [0, 3], \ J^1 = [3, 6], \ J^2 = [6, 9] \). For an illustration see Figure 3.1.

\[
A_0 = \begin{bmatrix} 1 & 0 & 0 \\ 6 & 3 & 3 \\ 1 & 3 & 3 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 3 & 1 & 0 \\ 3 & 6 & 3 \\ 0 & 1 & 3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & 3 & 1 \\ 3 & 3 & 6 \\ 0 & 0 & 1 \end{bmatrix}.
\]
Proof of Theorem 1.17. The statement follows from 1.10, as the Theorem is about the \( \text{proj}_{(1,0,0)} \)-projection (Example 2.1). An easy calculation yields that the spectral radius of \( p \cdot B_1 \) is \( 4 \cdot p \) which is less than 1 for \( p < 0.25 \). □

Proof of Theorem 1.18. Here we use what we wrote in Example 2.2. The first part follows from Theorem 1.9. The second part is an immediate consequence of Theorem 1.10. This is so, because the spectral radius of \( p \cdot A_0 \) in Example 2.2 is \( 6p \), which is less than 1 whenever \( p < \frac{1}{6} \). □

Proof of Theorem 1.19. The theorem is an easy consequence of Theorem 1.11 using again Example 2.2. □

Proof of Theorem 1.16. The proof is straightforward using Theorem 1.12. □

2.1. Proof of Theorem 1.14.

Lemma 2.3. We consider two independent copies of \( \mathcal{M}_p \), and we translate one of them by \((1,0,0)\). We call them \( \mathcal{M}^1_p \) and \( \mathcal{M}^2_p \). We denote their intersection by \( \mathcal{C} \). That is \( \mathcal{C} = \mathcal{M}^1_p \cap \mathcal{M}^2_p \) as in Figure 4. If \( p > \frac{1}{\sqrt{8}} = 0.35 \ldots \)

![Figure 4. Explanation of notations of Lemma 2.3.](image)

then there cannot exist a path from \( \mathcal{M}^1_p \) to \( \mathcal{M}^2_p \) through \( \mathcal{C} \), almost surely.

Proof. Assume that we have a path from \( \mathcal{M}^1_p \) to \( \mathcal{M}^2_p \) through \( \mathcal{C} \) with positive probability. Then for every \( n \), with positive probability, there exist a pair of retained level \( n \) cubes which are on opposite sides of \( \mathcal{C} \) and share a common face (contained in \( \mathcal{C} \)). Observe that the number of such level \( n \) pairs – denote it by \( \varepsilon_n \) – forms a Galton-Watson branching process with offspring-distribution \( \text{Bin}(8,p^2) \). It follows from the Theory of Branching
Processes (see [1, page 7]), that the process dies out almost surely in finite steps whenever
\[ 8 \cdot p^2 = \mathbb{E}(\varepsilon_n) < 1, \]
i.e. when \( p < \frac{1}{\sqrt{8}} \). Now if the process dies out in finite steps almost surely, then there cannot exists a connection between \( \mathcal{M}_p^1 \) and \( \mathcal{M}_p^2 \) through \( \mathcal{C} \). \( \square \)

**Proof of theorem 1.14.** Fix \( p_0 < \frac{1}{\sqrt{8}} \). Assume that the random Menger sponge \( \mathcal{M}_{p_0} \) contains a connected component with positive probability. Then with positive probability there exists an \( n \) such that with positive probability we can find two level \( n \) retained adjacent cube that are connected through their common face say \( D_1 \) and \( D_2 \). By statistical self-similarity, conditioned on \( D_1 \) and \( D_2 \) are two retained level \( n \) cubes, they are independent, rescaled copies of the random Menger sponge \( \mathcal{M}_p \). Hence, by Lemma 2.3 with probability one, \( D_1 \) and \( D_2 \) can not be connected through their common wall, which is a contradiction. \( \square \)

2.2. **The dual nature of \( \mathcal{M}_p \).** All until now we have always looked at \( \mathcal{M}_p \) as an example of a coin tossing self-similar set. However, to prove Theorem 1.15, we need another interpretation of \( \mathcal{M}_p \) as a three-dimensional inhomogeneous Mandelbrot percolation set. The random set what we call inhomogeneous Mandelbrot percolation in this paper simply named Mandelbrot percolation in [14], [16]. The motivation for this different interpretation of \( \mathcal{M}_p \) is that we would like to use the projection theorems of the paper [16], which are stated for the inhomogeneous Mandelbrot percolation sets. To construct \( \mathcal{M}_p \) as an inhomogeneous Mandelbrot percolation set, we divide the unit cube \( Q := [0, 1]^3 \) into 27 axes parallel cubes of side length \( 1/3 \).

\[
\{ \mathcal{R}_{(i,j,k)} \}_{i,j,k=0}^2, \text{ where } \mathcal{R}_{(i,j,k)} := \left[ \frac{i}{3}, \frac{i+1}{3} \right] \times \left[ \frac{j}{3}, \frac{j+1}{3} \right] \times \left[ \frac{k}{3}, \frac{k+1}{3} \right].
\]

Similarly, we define the level \( n \) cubes
\[
\mathcal{R}_n := \left\{ \mathcal{R}_{(i,j,k)} := (\bar{i}, \bar{j}, \bar{k}) + [0, 3^{-n}]^3 \right\}_{i,j,k \in \{0,1,2\}^n},
\]
where \( i := (i_1, \ldots, i_n) \) and \( \bar{i} := \sum_{\ell=1}^n i_\ell 3^{-\ell} \).

Let
\[
(2.3) \quad \tilde{A} := \{ (1, 1, 0), (1, 0, 1), (0, 1, 1), (1, 1, 2), (1, 2, 1), (2, 1, 1), (1, 1, 1) \}.
\]
Note that these indices corresponds to the cubes which are deleted when we construct the first approximation of the deterministic Menger sponge. For
each \((i, j, k) \in \{0, 1, 2\}^3\) we set a probability \(p_{(i,j,k)} \in [0,1]\) as follows:

\[
p_{(i,j,k)} := \begin{cases} 
0, & \text{if } (i,j,k) \in \tilde{A}; \\
p, & \text{otherwise}.
\end{cases}
\]

We retain the cube \(\mathcal{R}_{(i,j,k)}\) with probability \(p_{(i,j,k)}\) independently. Assume that we have already constructed the level \(n\) retained cubes. That is we are given the random set \(E_n \subset \{0, 1, 2\}^n \times \{0, 1, 2\}^n \times \{0, 1, 2\}^n\) such that after \(n\) steps we have retained the cubes \(\{\mathcal{R}_{(i,j,k)}\}_{(i,j,k) \in E_n}\). For every \((i, j, k) \in E_n\) we retain the cube \(\mathcal{R}_{(i_{n+1}, j_{n+1}, k_{n+1})}\) with probability \(p_{(i_{n+1}, j_{n+1}, k_{n+1})}\) independently of everything. Let \(E_n := \bigcup_{(i,j,k) \in E_n} \mathcal{R}_{(i,j,k)}\). Then \(\mathcal{M}_p = \bigcap_{n=1}^{\infty} E_n\).

**Proposition 2.4.** For any plane \(S(a, b, c)\):

\[
\frac{\text{Leb}_2(Q \setminus M_1 \cap S(a, b, c))}{\text{Leb}_2(Q \cap S(a, b, c))} \leq \frac{5}{9},
\]

where \(\text{Leb}_2\) denotes the two dimensional Lebesgue measure, \(M_1\) denotes the first approximation of the deterministic Menger sponge. Finally, \(S(a, b, c)\) denote the plane

\[
S(a, b, c) := \{(x, y, z) \in \mathbb{R}^3 : ax + by + c = z\}.
\]

2.3. **Proof of Theorem 1.15 assuming Proposition 2.4.** We use some of the results of Simon and Vágó [16] on projections of inhomogeneous Mandelbrot percolations. In [16] Simon and Vágó introduces conditions under which the projections of the inhomogeneous Mandelbrot percolation contains an interval almost surely conditioned on non-extinction. In what follows we show that for \(p > 0.25\) these conditions are satisfied. Let \(\Pi_{(a,b)}\) denote the projection along the planes \(S(a, b, \cdot)\) to the \(z\)-coordinate axis, namely

\[
\Pi_{(a,b)}(x, y, z) := z - ax - by.
\]

In particular, \(\Pi_{(a,b)}(S(a, b, t)) = t\). Observe that whenever \(\tilde{c} \neq 0\) we have

\[
\frac{1}{\tilde{c}} \text{proj}_{(\tilde{a}, \tilde{b}, \tilde{c})} \equiv \Pi_{(-\frac{a}{\tilde{c}}, -\frac{b}{\tilde{c}})}.
\]

We remark that because of the symmetries of the Menger sponge without loss of generality we can restrict our attention to those projections \(\text{proj}_{(-\tilde{a}, -\tilde{b}, \tilde{c})}\) for which \(\tilde{c} \neq 0\). These projections can be identified by projections along planes in the form of (2.6).

In particular, the following implication holds:

\[
\text{int}(\Pi_{(a,b)}(\mathcal{M}_p)) \neq \emptyset, \quad \forall (a, b) \implies \text{int}(\text{proj}_{(\tilde{a}, \tilde{b}, \tilde{c})}(\mathcal{M}_p)), \quad \forall (\tilde{a}, \tilde{b}, \tilde{c}),
\]
where \( \text{int} \) stands for the interior. Clearly, the inequality in (2.5) is equivalent to

\[
\frac{\text{Leb}_2(M_1 \cap S(a, b, c))}{\text{Leb}_2(Q \cap S(a, b, c))} \geq \frac{4}{9},
\]

(2.10)

**Proposition 2.5.** If the inequality in (2.10) holds for all \((a, b, c)\) then almost surely, for all \((a, b)\), the projection \(\Pi_{(a, b)}(M_p)\) contains an interval for all \(p > 0.25\).

**Proof.** Suggested by the formula on the top of page 177 of the paper [16] we consider the function,

\[
f(t) := \text{Leb}_2(S(a, b, t) \cap Q).
\]

(2.11)

A combination of [16, Theorem 1.2 and Proposition 2.3] applied to the projection \(\Pi_{(a, b)}\) and function \(f\) yields the assertion of the Proposition. More precisely, if we replace \(\text{proj}_\alpha\) in [16, Theorem 1.2] by \(\Pi_{(a, b)}\) then we obtain that the assertion of Proposition 2.5 holds if a certain condition which is named as Condition \(A(a, b)\) (see [16, Definition 2.1]) holds for all \((a, b)\). On the other hand, [16, Proposition 2.3] asserts in our case that another condition which is called Condition \(B(a, b)\) (see [16, Definition 2.2]) implies that Condition \(A(a, b)\) holds. It is immediate from the definitions that (2.10) implies that Condition \(B(a, b)\) holds with \(f\) defined in (2.11) and \(\Pi_{(a, b)}\). □

**Proof of Theorem 1.15 assuming Proposition 2.4.** The assertion of Theorem 1.15 immediately follows from Proposition 2.5 and the implication in (2.9). □

### 2.4. Proof of Proposition 2.4.

We often use the following simple observation.

**Fact 2.6.** Observe that \(Q \setminus M_1\) remains unchanged if we permute the coordinate axes and if we reflect to any of the planes \(x = \frac{1}{2}\), \(y = \frac{1}{2}\) or \(z = \frac{1}{2}\).

Hence, it is easy to see that without loss of generality we may assume, that

\[
0 \leq a \leq b \leq 1.
\]

(2.12)

Under this assumption, the plane \(S(a, b, c)\) intersects the unit cube \([0, 1]^3\) if and only if

\[
-(a + b) \leq c \leq 1.
\]

(2.13)

So, we can confine ourselves to the region (see Figure 5)
\begin{equation}
\mathcal{D} := \{(a, b, c) : 0 \leq a \leq b \leq 1, -(a + b) \leq c \leq 1\}.
\end{equation}

We partition \( \mathcal{D} = \mathcal{D}_{CA} \cup \mathcal{D}_{WM} \), where

\begin{equation}
\mathcal{D}_{WM} := \left\{(a, b, c) \in \mathcal{D} : a \in \left[\frac{1}{3}, 1\right], b \in [a, 1],
\right. \\
\left. c \in \left[\frac{2}{3} - (a + b), 0\right] \cup \left[\max\{0, 1 - (a + b)\}, \frac{1}{3}\right]\right\},
\end{equation}

and

\begin{equation}
\mathcal{D}_{CA} := \mathcal{D} \setminus \mathcal{D}_{WM}.
\end{equation}

According to this partition, the proof is divided into two major parts. We verify (2.5)

- for \((a, b, c) \in \mathcal{D}_{CA}\) with case analysis in Section 2.4.2.
- For \((a, b, c) \in \mathcal{D}_{WM}\) we do the following: We introduce a function \(\tilde{H}(a, b, c)\) on \(\mathcal{D}\) in (2.25), and we reformulate the inequality in (2.5) as \(\tilde{H}|_{\mathcal{D}_{WM}} \geq 0\). To verify this, in Section 2.4.3 we estimate the Lipschitz constant of \(\tilde{H}|_{\mathcal{D}_{WM}}\). In Section 2.4.4 we consider a sufficiently dense grid \(G \subset \mathcal{D}_{WM}\). Then using Wolfram Mathematica we calculate \(C := \min \tilde{H}|_G\). We verify that \(C > 0\) is so large that because of the Lipschitz property of \(\tilde{H}\), we get that \(\tilde{H}|_{\mathcal{D}_{WM}} > 0\).

First we introduce the notation used in the rest of Section 2.4.
2.4.1. Notations. We write \( \text{proj}_{x,y} \) for the orthogonal projection to the \((x, y)\)-coordinate plane, that is
\[
(2.17) \quad \text{proj}_{x,y} : \mathbb{R}^3 \to \mathbb{R}^2, \quad \text{proj}_{x,y}((u, v, w)) := (u, v).
\]
The area of the unit cube intersected with the plane \( S(a, b, c) \) and the area of the \( \text{proj}_{x,y} \)-projection respectively denoted by:
\[
(2.18) \quad F(a, b, c) := \text{Leb}_2(Q \cap S(a, b, c)), \quad \tilde{F}(a, b, c) := \text{Leb}_2(\text{proj}_{x,y}(Q \cap S(a, b, c))).
\]
It is easy to see that
\[
(2.19) \quad \tilde{F}(a, b, c) \sqrt{1 + a^2 + b^2} = F(a, b, c).
\]
Let
\[
(2.20) \quad A := \left\{ \left( \frac{0}{3}, \frac{1}{3}, \frac{1}{3} \right), \left( \frac{1}{3}, 0, \frac{1}{3} \right), \left( \frac{1}{3}, \frac{1}{3}, 0 \right), \left( \frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right), \left( \frac{2}{3}, \frac{1}{3}, \frac{1}{3} \right), \left( \frac{1}{3}, \frac{2}{3}, \frac{1}{3} \right) \right\}
\]
denote the lower left corners (the closest point of the cube to the origin) of the level-1 cubes missing from the first approximation of the Menger sponge. Also when we want to calculate the area of a level 1 square with lower left corner \((u, v, w)\) intersected with the plane \( S(a, b, c) \) we renormalize the plane with respect to the cube and calculate the area of the renormalized intersection and divide it by 9 (see (2.23)). Renormalization means that we transform a level 1 cube into the unit cube. In case of a level one cube \((u, v, w) + [0, \frac{1}{3}]^3\) this transformation is:
\[
(2.21) \quad (x, y, z) \to 3(x - u, y - v, z - w).
\]
The renormalization of the plane \( S(a, b, c) \) with respect to the level 1 cube with lower left corner \((u, v, w)\) is \( S(a, b, c') \), where \( c' := 3(au + bv + c - w) \). This is why we introduce
\[
(2.22) \quad g(a, b, c, u, v, w) := (a, b, 3(au + bv + c - w)).
\]
Observe that for a level 1 cube \( \hat{Q} := (u, v, w) + [0, \frac{1}{3}]^3 \):
\[
(2.23) \quad \text{Leb}_2(\hat{Q} \cap S(a, b, c)) = \frac{1}{9} F(g(a, b, c, u, v, w)).
\]
Hence, to verify Proposition 2.4 we need to show that for any \((a, b, c)\):

\[
(2.24) \quad H(a, b, c) := \frac{5}{9} \mathcal{L}eb_2(Q \cap S(a, b, c)) - \mathcal{L}eb_2(Q \setminus M_1 \cap S(a, b, c))
\]

\[
= \frac{5}{9} F(a, b, c) - \frac{1}{9} \sum_{(u,v,w) \in A} F(g(a, b, c, u, v, w)) \geq 0.
\]

Let

\[
(2.25) \quad \tilde{H}(a, b, c) := \frac{5}{9} \tilde{F}(a, b, c) - \frac{1}{9} \sum_{(u,v,w) \in A} \tilde{F}(g(a, b, c, u, v, w)).
\]

By (2.19) \(H(a, b, c) = \frac{1}{\sqrt{1+a^2+b^2}} \tilde{H}(a, b, c)\). In this way we have proved that

**Fact 2.7.** The inequality in (2.5) holds if

\[
(2.26) \quad \tilde{H}(a, b, c) \geq 0, \quad \text{for any } 0 \leq a \leq b \leq 1 \text{ and any } c.
\]

Let also

\[
(2.27) \quad \ell_i(x) := -\frac{a}{b}x + \frac{i}{3} - \frac{c}{b}, \quad i \in \{0, 1, 2, 3\}
\]

denote the line which is the \(\text{proj}_{x,y}\)-projection of

\[
\left\{ \left( x, y, \frac{i}{3} \right) : x \in \mathbb{R}, y \in \mathbb{R} \right\} \cap S(a, b, c).
\]

That is \(y = \ell_i(x)\) is the \(\frac{i}{3}\)-level set of \(S(a, b, c)\). Note that (2.12) implies that the slope of \(\ell_i(x)\) is between \(-1\) and \(0\).

For any \(H \subseteq [0, 1]^2\) the **bad part of** \(H\) is \(H \cap \text{proj}_{x,y}(S(a, b, c) \cap Q \setminus M_1)\) and the **good part of** \(H\) is \(H \cap \text{proj}_{x,y}(S(a, b, c) \cap M_1)\). The bad part of \([0,1]^2\) is going to be denoted by dark color on the figures below. Let \(Q_{i,j} := \left( \frac{i}{3}, \frac{j}{3} \right) + \left[0, \frac{1}{3}\right]^2\), \(i, j \in \{0, 1, 2\}\) as shown in the Figure 6A. We call these squares **level-1 squares**. By the construction of the Menger sponge, the squares in the corners can not have bad parts. The bad parts of \(Q_{1,1}\) is between the lines \(\ell_0(x)\) and \(\ell_3(x)\), and the bad parts of \(Q_{1,0}, Q_{2,1}, Q_{0,1}, Q_{1,2}\) are between the lines \(\ell_1(x)\) and \(\ell_2(x)\).

First we show by a detailed case analysis (Section 2.4.2) that for some values of \(a, b\) and \(c\) the assertion (2.26) holds. Then for the uncovered values of \(a, b, c\) we create a grid to calculate the value of \(\tilde{H}\) on the grid and estimate the value of the function out of the grid points, using the upper bound (Section 2.4.3) on the Lipschitz constant for the function \(\tilde{H}\).
2.4.2. Case analysis. Recall that we always assume that $0 \leq a \leq b \leq 1$. Note that if $a = b = 0$ and $c \in (\frac{1}{3}, \frac{2}{3})$ then $\tilde{H}(a, b, c) = 0$.

**Fact 2.8.** If $c > 0$ and $a + b + c \leq 1$, then $\tilde{H}(a, b, c) \geq 0$.

**Proof.** Whenever $c > 0$ and $a + b + c \leq 1$ by elementary geometry we have $\tilde{F}(a, b, c) = 1$. Hence, $\tilde{H}(a, b, c)$ is the smallest when all of $Q_{1,0}, Q_{2,1}, Q_{0,1}, Q_{1,2}$ is between the lines $\ell_1(x)$ and $\ell_2(x)$. In that case $\tilde{H}(a, b, c) = 0$. For visual explanation see Figures 6B and 6C. \hfill \Box

**Fact 2.9.** If $\frac{1}{3} \leq c \leq 1$, then $\tilde{H}(a, b, c) \geq 0$.

**Proof.** Whenever $a + b + c \leq 1$ we are done by Fact 2.8. Hence, we can assume, that $a + b + c > 1$. In this case $\ell_0([0, 1]), \ell_1([0, 1]) \leq 0$. This follows from $c \geq \frac{1}{3}$ and (2.27). For any such $S(a, b, c)$ plane if we fix the line $\ell_3(x)$ it is easy to see that the worst case scenario is when $c = \frac{1}{3}$, since for a fixed $\ell_3(x)$ the function $\tilde{F}(a, b, c)$ remains the same but the area of the bad part grows as we decrease $c$. Hence, we may assume, that $c = \frac{1}{3}$ and we know that $\ell_3(1) < 1$ (because $a + b + c > 1$) hence $a + b > \frac{2}{3}$. So, it completes the proof of Fact 2.9 if we verify that the following Claim holds:

**Claim 2.10.** If $a + b + c > 1$ and $c = \frac{1}{3}$ then for every level-1 square $Q_{i,j}$ we can find another level-1 square $Q_{i',j'}$ such that the area of the good part of $Q_{i',j'}$ is greater than or equal to the area of the bad part of $Q_{i,j}$. Moreover, distinct $(i', j')$ correspond to distinct $(i, j)$.

Namely, this Claim implies that the area of the bad part of the unit square is smaller than the area of its good part and hence, $\tilde{H}(a, b, c) > 0$. 

---

**Figure 6.** Figures for Fact 2.8 and 2.9. In the last three figures the lighter color denotes the good and the darker the bad part of the unit square respectively.
The proof of Claim 2.10 will be given below, but first we remark that this proof if illustrated on Figure 6d. The light background regions on Figure 6d refer to the good parts and the dark background regions refer to the bad parts. Figure 6d indicates that the bad region of every level-1 square is compensated by an identically patterned good region of a corresponding level 1 square having at least as large area as the corresponding bad region.

Proof of Claim 2.10.

- \(Q_{1,2}\) does not have a bad part in this case, because \(a + b > \frac{2}{3}\) and \(b > \frac{1}{3}\) hence \(a + 2b > 1\) thus \(\ell_2\left(\frac{1}{3}\right) = \frac{1-a}{3b} < \frac{2}{3}\). All of the lines \(\ell_i(x)\) has negative slope, hence this shows that all of the points of \(\ell_2(x)\) is under the square \(Q_{1,2}\).
- The bad part of \(Q_{1,1}\) is compensated by the good part of \(Q_{0,0}\), since \(\ell_3(0) \geq \ell_3\left(\frac{1}{3}\right) - \frac{1}{3}\) and hence the line segment \(\ell_3([0, \frac{1}{3}])\) is higher relative to the line \(y = 0\) than the segment \(\ell_3\left([\frac{1}{3}, \frac{2}{3}]\right)\) to the line \(y = \frac{1}{3}\). This part is illustrated by green lines on Figure 6d.
- The area of the bad part of \(Q_{1,0}\) is smaller than the area of the good part of \(Q_{2,0}\) since \(\ell_3\left(\frac{2}{3}\right) \geq \ell_2\left(\frac{1}{3}\right)\) and \(\ell_2(x)\) and \(\ell_3(x)\) has the same slope. This part is illustrated by pink honeycombs on Figure 6d.
- The area of the bad part of \(Q_{2,1}\) is at most as much as the area of the good part of \(Q_{2,2}\). This follows from the fact that \(\ell_3\left(\frac{2}{3}\right) - \frac{2}{3} \geq \ell_2\left(\frac{1}{3}\right) - \frac{1}{3}\). This part is illustrated by orange lines on Figure 6d.
- The area of the bad part of \(Q_{0,1}\) is at most as much as the area of the good part of \(Q_{0,2}\). This follows from the fact that \(\ell_3(0) - \frac{2}{3} \geq \ell_2(0) - \frac{1}{3}\). This part is illustrated by blue lines (crossed) on Figure 6d.

As we mentioned above, Claim 2.10 implies that the assertion of Fact 2.9 holds.

Fact 2.11. If \(a + b + c \leq \frac{2}{3}\), then \(\tilde{H}(a, b, c) \geq 0\).

Proof. This is immediate from Facts 2.6 and 2.9.

Corollary 2.12. If \(a + b \leq \frac{2}{3}\), then \(\tilde{H}(a, b, c) \geq 0\).

Proof. Whenever

- \(0 < c \leq \frac{1}{3}\): we are done by Fact 2.8;
- \(\frac{1}{3} \leq c \leq 1\): , then
  \(a + b < 1 - c\): we are done by Fact 2.8,
\( a + b \geq 1 - c \): we are done by Fact 2.9;

c \leq 0: we are done by Fact 2.11.

\[ \square \]

**Fact 2.13.** If \( a < \frac{1}{3} \), then \( \tilde{H}(a, b, c) \geq 0 \).

Proof. Let \( h_i := \ell_i(0) \) and \( \tilde{h}_i := \ell_i(1) \). Since \( \text{dist}(h_i, h_{i+1}) = \frac{1}{36} \) and \( b \leq 1 \) it follows, that \( \text{dist}(h_i, h_{i+1}) \geq \frac{1}{3} \), hence the following cases are possible:

I. \( h_0 \in \left[ 0, \frac{1}{3} \right] \)
   I.a. \( h_1 \in \left[ \frac{1}{3}, \frac{2}{3} \right] \)
   I.a.1. \( h_2 \in \left[ \frac{2}{3}, 1 \right] \)
   I.a.2. \( h_2 \geq 1 \)
   I.b. \( h_1 \in \left[ \frac{2}{3}, 1 \right] \) and \( h_2 \geq 1 \)
   I.c. \( h_1 \geq 1 \) and \( h_2 > 1 \)

II. \( h_0 \in \left[ \frac{1}{3}, \frac{2}{3} \right] \)
   II.a. \( h_1 \in \left[ \frac{2}{3}, 1 \right] \) and \( h_2 \geq 1 \)
   II.b. \( h_1 \geq 1 \) and \( h_2 > 1 \)

III. \( h_0 \in \left[ \frac{2}{3}, 1 \right] \) and \( h_1 \geq 1 \) and \( h_2 > 1 \)

IV. \( h_0 \geq 1 \)

V. \( h_0 < 0 \).

The proof of Case V. follows from Fact 2.6, Fact 2.8 and from Cases I-IV. We mainly use two methods for the proof.

One is to estimate the area in the case when \( a = 0 \) i.e. the slope of the lines \( \ell_i \) equals zero, then we upper bound the growth of the area of the bad part and lower bound the shrinkage of the area of the good part to get a "worst case scenario" fraction of the bad and the total area. We use this method in case of I.a.1.-I.b..

The other way is to prove that the good part of the unit square is at least as big as the bad part. We do this one-by-one, for each \( Q_{i,j} \) that has bad part. We apply this method in case of I.c.-IV..

**Case I.a.1.:** In this case \( h_0 \in \left[ 0, \frac{1}{3} \right] \), \( h_1 \in \left[ \frac{1}{3}, \frac{2}{3} \right] \), \( h_2 \in \left[ \frac{2}{3}, 1 \right] \) and \( h_3 \geq 1 \). We distinguish two cases: when \( a = 0 \) or \( a > 0 \).

\( a = 0 \): Then \( \ell_i(x) = h_i \) and \( \tilde{F}(a, b, c) = 1 - h_0 \). \( Q_{1,1} \) has only bad part since \( h_0 \leq \frac{1}{3} \) and \( h_1 \geq 1 \), \( Q_{1,0} \) has no bad parts because \( h_1 \geq \frac{1}{3} \). The area of the bad part of \( Q_{0,1} \) and \( Q_{2,1} \) together is \( \frac{2}{3} \left( \frac{2}{3} - h_1 \right) \) and the area of the bad part of \( Q_{1,2} \) is \( \frac{1}{3} \left( h_2 - \frac{2}{3} \right) \). So, the bad area in \([0, 1]^2 \) is \( 1/3 \) of the total area which is \( \tilde{F}(a, b, c) = 1 - h_0 \).
\(a > 0\): In this case we distinguish two further sub-cases. Namely, the first is when \(\ell_3(1) \geq 1\) (see Figure 7A) and the second one is when \(\ell_3(1) < 1\) (see Figure 7B). In both sub-cases the area of the bad part increases at most with the area of the triangle \((0, h_1), (1, h_1), (1, \hat{h}_1)\) intersected with \(Q_{0,1}\) and \(Q_{1,0}\) and \(Q_{2,1}\) (see the light hatched area on Figure 7B). This area is smaller than \(\frac{a}{3b}\).

\(\ell_3(1) \geq 1\): Then the total area (the area of the good and the bad part together) does not decrease. Rewriting \(h_2 = h_0 + \frac{2}{3b}\) and \(h_1 = h_0 + \frac{1}{3b}\) gives that \(\tilde{H}(a, b, c) \geq \frac{2}{9}(1-h_0) - \frac{a}{3b}\). It follows from \(h_2 = \frac{2}{3b} - c < 1\), that \(b + c > \frac{2}{3}\) and from \(\frac{2}{3b} = h_2 - h_0 \leq 1\), that \(b > \frac{2}{3}\). Multiplying \(\frac{2}{9}(1-h_0) - \frac{a}{3b}\) by \(9b > 0\) gives \(2b + c - 3a > \frac{1}{3} > 0\). Thus, \(\tilde{H}(a, b, c) > 0\).

\(\ell_3(1) < 1\): Then the total area decreases. An easy calculation shows that the total area is \(1 - \frac{c^2}{2ab} - \frac{(a+b+c-1)^2}{2ab}\). Observe that \(\ell_3(1) = -\frac{a}{b} + \frac{1-c}{b} < 1\) implies that \(a + b + c > 1\). Using this and that \(b > \frac{2}{3}\) a substantial but elementary calculation shows that in this case \(\tilde{H}(a, b, c) > 0\).

**Case I.a.2.:** In this case \(\hat{h}_3 > 1\) since \(\hat{h}_3 > h_2 > 1\). The proof is very similar to the one of I.a.1..

\(a = 0\): In this case the total area \(\tilde{F}(a, b, c)\) is \(1 - h_0\). From \(h_1 \geq \frac{1}{3}\) it follows that \(Q_{1,0}\) does not have a bad part. Hence the bad parts and their areas are as follows:

- \(Q_{1,1}\) and \(Q_{1,2}\) consists only of bad parts, hence their bad area together is \(\frac{2}{3}\),

- \(Q_{0,1}\) and \(Q_{2,1}\) have bad parts over \(h_1\), the area of their bad parts together is \(\left(\frac{2}{3} - h_1\right) \cdot \frac{2}{3}\)

Altogether the bad part has area

\[
\frac{2}{3} (1 - h_0) - \frac{2}{9b}.
\]

\(2b(1-h_0) < 2\), hence \(6b(1-h_0) < 2 + 4b(1-h_0)\) thus \(\frac{2}{3} - \frac{2}{9b(1-h_0)} < \frac{4}{9}\), hence the fraction of the bad area and the total is smaller than \(\frac{4}{9} < \frac{5}{9}\).

\(a > 0\): When \(a > 0\), it is clear that the area bad area grows with at most \(\frac{a}{3b}\) similarly to the previous case. The total area does not decrease, hence we can use the previous calculation to show that in this case the assertion holds.

**Case I.b.:**
a = 0: The total area is again \((1 - h_0)\). It follows from \(h_1 > \frac{2}{3}\), that only \(Q_{1,2}\) and \(Q_{1,1}\) can have bad parts. It is easy to see, that \(Q_{1,1}\) has only bad parts, and that the bad part of \(Q_{1,2}\) has area \((1 - h_1)\frac{1}{3} = \frac{1}{3} \left(1 - h_0 - \frac{1}{3}b\right)\). An easy calculation shows that in this case the fraction of the bad and the total area is smaller than \(\frac{4}{9}\), hence we are done.

\(a > 0\): Similarly to the earlier cases we consider the growth of the area if we increase \(a\). The additional bad part is the intersection of \(Q_{0,1} \cup Q_{2,1} \cup Q_{1,2}\) with the triangle which we get by intersecting the unit square with the area between \(\ell_1\) and the line with height \(h_1\). The area of this triangle is smaller than \(\frac{a}{3b}\). The assertion follows if we rewrite the inequality \(\frac{1}{9(1-h_0)} + \frac{1}{3} + \frac{3a-1}{9b(1-h_0)} < \frac{5}{9}\), getting \(1 + \frac{3a-1}{b} < 2(1-h_0)\), which trivially holds, since \(3a - 1 < 0\), hence the left hand side is smaller than 1, and the right hand side is greater than 1 by \(1 - h_0 > \frac{2}{3}\).

Case I.c.: This case is almost identical to II.b., only in that case the total area might be smaller because \(h_0\) is greater. For this reason instead of proving this statement we present the proof for Case II.b. later because that might seem more complicated.

Case II.a. Observe that from \(\ell_1(1) \geq h_0 \geq \frac{1}{3}\) it follows that \(Q_{1,0}\) does not have bad part.

\(Q_{1,2}\): The bad parts of it are compensated by the good parts of \(Q_{0,2}\). This is because \(Q_{0,2}\) only has good parts, because \(\ell_3(x) \cap [0,1] > 1\) and \(\ell_0(0) \leq \frac{2}{3}\).

\(Q_{1,1}\): \(Q_{2,2}\) only has good parts, this can be verified in the same way as in the case of \(Q_{0,2}\). Thus the bad part of \(Q_{1,1}\) is compensated by \(Q_{2,2}\).

\(Q_{0,1}\): Since \(b \leq 1\), \(\ell_1(x) - \ell_0(x) = \frac{1}{3b} \geq \frac{1}{3}\). Consequently the area of \(Q_{0,0}\) above the line \(\ell_0(x)\) (i.e. the good part of \(Q_{0,0}\)) is greater than the part of \(Q_{0,1}\) above \(\ell_1(x)\) (i.e. the bad part of \(Q_{0,1}\)). It follows, that the area of the bad part of \(Q_{0,1}\) is smaller than the area of the good part of \(Q_{0,0}\).

\(Q_{2,1}\): By almost exactly the same reasoning as in the previous case the area of the bad part of \(Q_{2,1}\) is smaller than the good part of \(Q_{2,1}\).

Case II.b.: In this case \(h_0 \in \left[\frac{1}{3}, \frac{2}{3}\right]\), \(h_1 \geq 1, h_2, h_3 > 1\). Note that \(a < \frac{1}{3}\) yields \(\ell_3([0,1]) > h_1 \geq 1\) and since \(h_0 < \frac{2}{3}\) the squares \(Q_{0,1}\) and \(Q_{2,2}\) has only good parts. It follows from \(a < \frac{1}{3}\) that \(\ell_1([0,1]) > h_0 \geq \frac{1}{3}\) and \(h_1 \geq 1\). Potentially only \(Q_{1,1}, Q_{1,2}\) and \(Q_{2,1}\) have bad parts. For the illustration of...
the following pairing see Figure 7C, where the bad parts are denoted by the darker, the good parts are denoted by the lighter color, the pairs has the same pattern and color.

**Q**_{1,1}: Since the area of the bad part of Q_{1,1} is at most $\frac{1}{9}$, and the area of Q_{2,2} is $\frac{1}{9}$ the area of the bad part of Q_{1,1} is compensated by the area of the good part of Q_{2,2}.

**Q**_{1,2}: The same holds for the bad area of Q_{1,2} and the good area of Q_{0,2}.

**Q**_{2,1}: Q_{2,1} has bad part only in the case, when $\hat{h}_1 < \frac{2}{3}$. The good part of Q_{2,0} which is above $\ell_0(x)$ and the bad part of Q_{2,1} which is above $\ell_1(x)$ and since $\ell_1(x)$ and $\ell_0(x)$ has the same slope, it is enough to show that $\frac{1}{3} - \hat{h}_0 > \frac{2}{3} - \hat{h}_1$. Since $\hat{h}_1 = \hat{h}_0 + \frac{1}{3b}$ the inequality reduces to $1 > b$, which is true by our assumptions, hence the area of the good part of Q_{2,0} is greater than the area of the bad part of Q_{2,1}.

**Case III.** It is clear that only Q_{1,1} and Q_{1,2} can have bad parts, since $\ell_1 \cap [0, 1] \geq h_0 \geq \frac{2}{3}$.

**Q**_{1,1}: Whenever Q_{1,1} has bad parts, $\ell_0(\frac{2}{3}) \leq \frac{2}{3}$, then the whole Q_{2,2} is good. Thus it compensates for the bad parts of Q_{1,1}.

**Q**_{1,2}: Since $\frac{1-a}{3b} > 0$ it is easy to see that the area of the good part of Q_{0,2} is greater than the area of the bad part of Q_{1,2}.

**Case IV.** In this case only Q_{1,1} can have bad part, and if it happens then the whole Q_{2,2} is good, hence the bad area of Q_{1,1} is compensated.

![Figure 7](image)

**Figure 7.** Figures for Fact 2.13. In the first two figures the red lines denotes the height of $h_i$, the blue lines denotes $\ell_i(x)$. □
2.4.3. Lipschitz-constants for the function $\tilde{F}(a, b, c)$. Putting together Facts 2.13 and 2.7 from now we may always assume that

$$\frac{1}{3} \leq a \leq b \leq 1.$$  

Since the gradient of the renormalized plane remains the same these assumptions holds when instead of $\tilde{F}$ we consider $\tilde{F} \circ g$ ($g$ was defined in (2.22)). For a fix $(a, b)$ satisfying (2.28) consider the plane $S(a, b, c)$. This plane can intersect the unit cube $[0, 1]^3$ in eight different ways depending on $c$. The different ways of intersections yields different formulas for $\tilde{F}(a, b, c)$.

We estimate the Lipschitz constant from above by the sup-norm of the gradient of $\tilde{F}$, which can be calculated using elementary calculus. Table 1 contains the results of the calculation, and some figures to visualize the different cases.

\[
\begin{array}{cccccccc}
 a + b < 1 \\
-1 & -(a + b) & -b & -a & 0 & 1 - (a + b) & 1 - b & 1 - a & 1 \\
\text{case 2} & \text{case 3} & \text{case 4} & \text{case 1} & \text{case 6} & \text{case 7} & \text{case 8} & \\
 a + b > 1 \\
-(a + b) & -1 & -b & -a & 1 - (a + b) & 0 & 1 - b & 1 - a & 1 \\
\text{case 2} & \text{case 3} & \text{case 4} & \text{case 5} & \text{case 6} & \text{case 7} & \text{case 8} & \\
\end{array}
\]

**Figure 8.** Figure for Table 1. The two figures show that for a fixed $a$ and $b$ what case does the different values of $c$ give. The cases are named after the first column of Table 1, e.g. when $-b \leq c \leq -a$ we are in case 3. In the upper figure we consider the case, when $a + b < 1$, in the lower, when $a + b > 1$.

Since according to the Table all the Lipschitz-constants are less then 4,

$$\| \tilde{F}(a, b, c) - \tilde{F}(\hat{a}, \hat{b}, \hat{c}) \| < 4 \| (a, b, c) - (\hat{a}, \hat{b}, \hat{c}) \|.$$  

Using that

$$A := \nabla_{a, b, c} g(a, b, c, u, v, w) = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
3u & 3v & 3
\end{bmatrix}.$$  

| Name | Figure 1 | Figure 2 | Conditions | $\tilde{F}(a, b, c)$ | Lipsch. |
|------|----------|----------|------------|-----------------|--------|
| 0    |          |          | $c \geq 1$  | 0               | 0      |
|      |          |          | or $a + b + c \leq 0$ |                 |        |
| 1    |          |          | $c \geq 0$  | 1               | 0      |
|      |          |          | and $a + b + c \leq 1$ |                 |        |
| 2    |          |          | $-(a + b) \leq c \leq b$ | $\frac{(a+b+c)^2}{2ab}$ | $\frac{3\sqrt{3}}{2}$ |
| 3    |          |          | $-b \leq c \leq -a$ | $\frac{a+2b+2c}{2b}$ | $\frac{3\sqrt{3}}{2}$ |
| 4    |          |          | $-a \leq c$  | $1 - \frac{c^2}{2ab}$ | $\frac{3\sqrt{3}}{2}$ |
|      |          |          | and $c \leq \min\{0, 1 - (a + b)\}$ |                 |        |
| 5    |          |          | $1 \leq a + b$ | $1 - \frac{c^2+(a+b+c-1)^2}{2ab}$ | $\frac{131}{54}$ |
|      |          |          | and $1 - (a + b) \leq c \leq 0$ |                 |        |
| 6    |          |          | $\max\{0, 1 - (a + b)\} \leq c$ | $1 - \frac{(a+b+c-1)^2}{2ab}$ | $\frac{3\sqrt{3}}{2}$ |
|      |          |          | and $c \leq 1 - b$ |                 |        |
| 7    |          |          | $1 - b \leq c \leq 1 - a$ | $\frac{2-2c-a}{2b}$ | $\frac{3\sqrt{3}}{2}$ |
| 8    |          |          | $1 - a \leq c \leq 1$ | $\frac{(1-c)^2}{2ab}$ | $\frac{3\sqrt{3}}{2}$ |

**Table 1.** The first column contains the name (used in Figure 8) of the case. The second column shows the way of the intersection of the plane and the unit cube, the third column shows the projection of the above. In the fourth – conditions – column, we describe the region of $(a, b, c)$ under which we are in the given case. The fifth column contains the value of the function $\tilde{F}$ in the given case, and the last, Lipsch., column contains the upper bound for the Lipschitz constant in the given case.
and that $\sup_{(u,v,w)\in A} \|A\| < 4$, it follows, that

\begin{equation}
(2.31) \quad \left\| \sum_{(u,v,w)\in A} \tilde{F}(g(a, b, c, u, v, w)) - \tilde{F}(g(\hat{a}, \hat{b}, \hat{c}, u, v, w)) \right\| < \sum_{(u,v,w)\in A} 4 \cdot \left\| g(a, b, c, u, v, w) - g(\hat{a}, \hat{b}, \hat{c}, u, v, w) \right\|
\leq |A| \cdot 4 \cdot 4 \left\| (a, b, c) - (\hat{a}, \hat{b}, \hat{c}) \right\| \leq 7 \cdot 4 \cdot 4 \left\| (a, b, c) - (\hat{a}, \hat{b}, \hat{c}) \right\|.
\end{equation}

\begin{equation}
(2.32) \quad \|\tilde{H}(a, b, c) - \tilde{H}(\hat{a}, \hat{b}, \hat{c})\| \leq \frac{5}{9} \|\tilde{F}(a, b, c) - \tilde{F}(\hat{a}, \hat{b}, \hat{c})\|
+ \frac{1}{9} \left\| \sum_{(u,v,w)\in A} \tilde{F}(g(a, b, c, u, v, w)) - \tilde{F}(g(\hat{a}, \hat{b}, \hat{c}, u, v, w)) \right\|
< \frac{44}{3} \left\| (a, b, c) - (\hat{a}, \hat{b}, \hat{c}) \right\| < 15 \left\| (a, b, c) - (\hat{a}, \hat{b}, \hat{c}) \right\|,
\end{equation}

where the second inequality follows from the combination of (2.29) and (2.31). Hence,

\begin{equation}
(2.33) \quad \tilde{H}(\hat{a}, \hat{b}, \hat{c}) > \tilde{H}(a, b, c) - 15 \left\| (a, b, c) - (\hat{a}, \hat{b}, \hat{c}) \right\|.
\end{equation}

2.4.4. Numerical calculations using Wolfram Mathematica. First we construct a grid $G$ depending on a given positive number $\hat{d}$. Let $a_0 = \frac{1}{3}$ and if $a_n$ is defined and $a_n + \hat{d} \leq 1$, let $a_{n+1} = a_n + \hat{d}$, otherwise $I = n$. For a given $i \in \{0, \ldots, I\}$, let $b_{i,0} = a_i$ and if $b_{i,n} + \hat{d} \leq 1$ then we define $b_{i,n+1} = b_{i,n} + \hat{d}$, otherwise $J_i = n$. For a given $i \in \{0, \ldots, I\}$, let $c_{i,j,0} = \frac{2}{3} - (a + b) \leq 0$ and if $c_{i,j,n}$ is defined and $c_{i,j,n} + \hat{d} \leq \frac{2}{3}$ then we put $c_{i,j,n+1} = c_{i,j,n} + \hat{d}$ otherwise $K_{i,j} = n$. Now if we choose a point $(\hat{a}, \hat{b}, \hat{c})$ such that $\frac{1}{3} \leq \hat{a} \leq 1$, $\hat{a} \leq \hat{b} \leq 1$ and $\frac{2}{3} - (a + \hat{b}) \leq \hat{c} \leq \frac{1}{3}$, then there exists a grid point $(a_i, b_{i,j}, c_{i,j,k}) \in G$, $i \in \{0 \ldots I\}$, $j \in \{0 \ldots, J_i\}$, $k \in \{0 \ldots, K_{i,j}\}$ such that $\| (a_i, b_{i,j}, c_{i,j,k}) - (a_i, b_{i,j}, c_{i,j,k}) \| < \sqrt{3} \hat{d}$. Observe that (2.33), yields the following implication: If $\tilde{H}(a, b, c) \geq 15 \cdot \sqrt{3} \cdot \hat{d}$ holds for every $(a, b, c) \in G$ then we have $\tilde{H}(a, b, c) \geq 0$ for every $(a, b, c)$. We choose $\hat{d} = \frac{1}{500}$ and with Wolfram Mathematica we evaluate $\tilde{H}$ at every grid point. We obtain that the minimum of the results is $\frac{62509}{1125000}$. This is greater than $15 \cdot \sqrt{3} \cdot \frac{1}{500}$.

3. Coin-tossing integer self-similar sets on the line

We consider the IFS $\mathcal{F} = \{f_i\}_{i=0}^{M-1}$ defined in Example 1.5. We would like to invoke the notation introduced in Definition 1.8. To do so without loss
of generality we may assume that the distinct elements of \( \mathcal{F} \) are the first \( m \) elements of \( \mathcal{F} \). Using them we form a new IFS

\[
(3.1) \quad \mathcal{S} := \left\{ S_i(x) := \frac{1}{L} x + t_i \right\}_{i=0}^{m-1}.
\]

Moreover, also without loss of generality we may assume, that

\[
(3.2) \quad L \geq 2, \ L \in \mathbb{N}, \text{ and } t_0, \ldots, t_m \in \mathbb{Z}, \ 0 = t_0 < t_1 < \cdots < t_{m-1}, \ L - 1|t_{m-1}.
\]

Recall from 1.12 in Definition 1.8 that \( n_j = \# \{ f_i \in \mathcal{F} : f_i = S_j \} \). From now on we denote the natural number (see (3.2)) \( \tilde{n} := \frac{t_m - 1}{L - 1} \). In this case the interval

\[
(3.3) \quad I := [0, \tilde{n} \cdot L]
\]
satisfies (1.3).

**Lemma 3.1** ([15]).

1. For all \( \ell \in [M], k \in [N] \) there exist \( i \in [N] \) and \( b \in [L] \) such that \( S_\ell(J^k) = J^i_b \).
2. \( S_\ell(J^k) = J^i_b \) if and only if \( S_\ell(J^k_a) = J^i_{ba} \) for all \( a \in [L]^k \).
3. If \( S_\ell^{-1}(J^i_{ba}) \neq J^k_a \) holds for all \( k \in [N] \), then \( \nu(S_\ell^{-1}(J^i_{ba})) = 0 \).

**Corollary 3.2.** For all \( \ell \in [m]^n \) and \( k \in [N] \) there exist an \( i \in [N] \) and \( a \in [L]^n \) such that

\[
S_\ell(J^k) = J^i_a.
\]

**Proof.** It follows from the first and second part of the previous Lemma (3.1) and mathematical induction.

It is easy to see, that the matrices are well-defined. Namely, for a given \( \ell \) and \( k \), for distinct \( i_1 \) and \( i_2 \) satisfying \( S_{i_1}(J^k) = J^\ell_a = S_{i_2}(J^k) \), we have \( t_{i_1} = t_{i_2} \), which is a contradiction, since by definition the translations differ.

Recall that the matrices \( \{ A_a \}_{a \in [L]} \) were defined in (1.18). Let

\[
(3.4) \quad A := \sum_{a \in [L]} A_a.
\]

The meaning of \( A_a(\ell, k) \) is simply the number of indices \( (i_1, \ldots, i_n) \in [M]^n \) such that \( f_{i_1,\ldots,i_n}(J^k) = J^\ell_a \), which is stated in the following lemma.

**Lemma 3.3.** For any \( n \in \mathbb{N}, \ n \geq 1, \ a \in [L]^n \) and \( \ell, k \in [N] \)

\[
(3.5) \quad A_a(\ell, k) = \# \{ (i_1, \ldots, i_n) \in [M]^n : f_{i_1,\ldots,i_n}(J^k) = J^\ell_a \}
\]
Proof. Follows easily from mathematical induction on \( n \). □

In what follows we denote

\[
\|A\| := e^T \cdot A \cdot e, \quad e^T = (1, \ldots, 1) \in \mathbb{R}^N.
\]

Note that we have non-negative matrices, hence the norm defined above in our case coincide with \( \|A\|_{1,1} \).

3.1. The probability space. First of all we define coin tossing self-similar sets precisely. Throughout this section we follow the method of Falconer and Xiong [9], only we simplify it a little, since the construction we use is way more simple. First let \( \hat{\Omega} := \{0, 1\}^M \), be the subsets of \([M]\), and let \( \hat{\mathcal{A}} \) be the discrete sigma-algebra on it. For an \( \omega = (\omega_1, \ldots, \omega_M) \in \hat{\Omega} \) and \( k = \# \{ \ell : \omega_\ell = 1 \} \):

\[
\hat{\mathbb{P}}(\{(\omega_1, \ldots, \omega_M)\}) := p^k(1 - p)^{M-k}.
\]

It is easy to see that \( \hat{\mathbb{P}} \) is a probability measure on the measurable-space \((\hat{\Omega}, \hat{\mathcal{A}})\). On this space we define the random variable

\[
X(\omega) := (X_1(\omega), \ldots, X_M(\omega)), \text{ where } X_k(\omega) = \omega_k.
\]

For the M-ary tree \( T \), we define

\[
(\Omega, \mathcal{A}, \mathbb{P}) := \bigotimes_{i \in T} (\Omega_i, \mathcal{F}_i, \mathcal{P}_i), \text{ where } (\Omega_i, \mathcal{A}_i, \mathcal{P}_i) = (\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mathbb{P}}).
\]

For each \( i \in T \), we define the

\[
(3.10) \quad \pi_i : \Omega \to \Omega_i, \quad \pi_i(\omega) := \omega^i.
\]

We also define

\[
(3.11) \quad X^{[i]} := X \circ \pi_i, \text{ i.e. } X^{[i]}_j = \begin{cases} 1, & \text{if } \omega^i_j = 1; \\ 0, & \text{otherwise}. \end{cases}
\]

Hence, \( X^{[i]} \) are i.i.d random variables with the same distribution as of \( X \). For an \( i = (i_1, \ldots, i_n) \in T \), we define

\[
(3.12) \quad X^i(\omega) := X^{[i_1]}(\omega) \cdots X^{[i_n]}(\omega),
\]

and the level \( n \)-th and the eventual survival set:

\[
(3.13) \quad \mathcal{E}_n(\omega) := \{ i \in [M]^n : X^i(\omega) = 1 \}, \quad \mathcal{E}_\infty(\omega) := \{ i \in \Sigma^M : i|_n \in \mathcal{E}_n(\omega) \forall n \in \mathbb{N} \}.
\]
We are given the deterministic self-similar IFS $F$ and $S$ as in Definition 1.8 on the line. Put

\begin{equation}
E_n(\omega) := \bigcup_{i \in \mathcal{E}_n(\omega)} I_i,
\end{equation}

where $I_i := f_i(I)$ (recall, that $I$ was defined in (3.3)). The Coin tossing integer self-similar set on the line corresponding to probability $p$ and the IFS $F$ is defined as

\begin{equation}
\Lambda_F(p) = \Lambda_F(p, \omega) := \bigcap_{n=1}^{\infty} E_n(\omega).
\end{equation}

3.2. **Proof of Theorem 1.9.** The proof uses the method introduced by Dekking and Simon [6].

We say that the interval $J$,

- is of type $V$ if $J = f_i(J^V)$ for some $i \in \mathcal{E}_n$ for some $n$;
- is of type $V$ with multiplicity $s \geq 0$ if

\[
\# \{i \in \mathcal{E}_n : J = f_i(J^V)\} = s.
\]

For $n \geq 0$; $U, V \in [N]$ and $a \in [L]^n$, let

\[
S_{U,V}(a) := \{i \in [M]^n : i \in \mathcal{E}_n \text{ and } f_i(J^V) = J^U_a\}.
\]

Thus, $J^U_a$ is of type $V$ with multiplicity $\#S_{U,V}(a)$. For a $V \in [N]$ let \(\{S_{U,V}(a)\}_{U \in [N]}\) be independent random variables such that the distribution of $S_{U,V}(a)$ is equal to the distribution of $S_{U,V}(a)$, and let

\[
S^V(a) := \bigcup_{U \in [N]} S_{U,V}(a).
\]

The following lemma illuminates the meaning of the previously introduced random variables.

**Lemma 3.4.** For any $U, V \in [N]$ and $a \in [L]^n$:

\begin{equation}
\mathbb{E}(\#S_{U,V}(a)) = p^n A_{a}(U, V)
\end{equation}

and

\begin{equation}
\mathbb{E}(\#S^V(a)) = p^n CS_{a,V}
\end{equation}
Proof. By Lemma 3.3 in the deterministic case the first part follows with $p = 1$, and for every level $n$ cylinder the probability of retention is $p^n$, hence the assertion follows. For the second part:

\[ (3.18) \quad \mathbb{E}(\#S^V(a)) = \sum_{U \in [N]} \mathbb{E}(\#S^{U,V}(a)) = p^n \sum_{U \in [N]} A_a(U, V) \]

Lemma 3.5. There exists a $U \in [N]$ and $a \in [L]^n$ such that

\[ (3.19) \quad \mathbb{P}\left( \{ \forall V \in [N] : \#S^{U,V}(a) > 0 \} \right) > 0 \]

Proof. The lemma follows from the second condition of the theorem, namely: under that condition there exist a $U \in [N]$ and $a \in [L]^n$ for some $n$ such that $A_a(U, V) > 0$ for all $V \in [N]$. The events $\{ \#S^{U,V}(a) > 0 \}_{V=0}^{N-1}$ are not exclusive and each has positive probability, hence the probability that all of them happens simultaneously is also positive. □

Fix $\tilde{U}$ and $\tilde{b}_k \in [L]^k$ in a way that lemma 3.5 holds for $U = \tilde{U}$ and $a = \tilde{b}_k$. Let

\[ (3.20) \quad \gamma_V := p \cdot \min_{a \in [L]} CS_{a,V} \quad \text{and} \quad \gamma := \min_{V \in [N]} \gamma_V. \]

Note that $\gamma > 1$ is equivalent to the the first condition of Theorem 1.9. That is

\[ (3.21) \quad \gamma > 1 \iff p \cdot CS_{a,j} > 1 \quad \text{for all} \quad a \in [L] \quad \text{and} \quad j \in [N]. \]

Let $H^0$ be the event in Lemma 3.5: that the interval $J^{\tilde{U}}_{\tilde{b}_k}$ is of every type, i.e.

\[ (3.22) \quad H^0 := \left\{ \forall V \in [N] : S^{\tilde{U},V}(\tilde{b}_k) > 0 \right\}. \]

Then by Lemma 3.5,

\[ (3.23) \quad p_0 := \mathbb{P}(H^0) > 0. \]

This means that with positive probability we can find

\[ (3.24) \quad i^{0,0}, \ldots, i^{0,N-1} \in [M]^k \quad \text{such that} \quad f_{i^{0,V}}(J^V) = J^{\tilde{U}}_{\tilde{b}_k}, \]

meaning that we can find indices that make $J^{\tilde{U}}_{\tilde{b}_k}$ of every type.

In what follows we define several processes counting the multiplicity of different types of the intervals of the form $J_{a_n}^V$ for $a_n = a_1, \ldots, a_n \in [L]^n$ and $V \in [N]$ for the different values of $n \in \mathbb{N}$. We apply large deviation theory to prove that these processes simultaneously do not die out with positive
probability, implying that the random attractor contains an interval with positive probability. Lastly a standard argument reveals that this is a 0-1 event conditioned on non-extinction.

We say that $i \in [M]^*$ makes $J_{a}^{U} (U \in [N], a \in [L^{[i]}])$ a type $V (V \in [N])$ interval if the following two holds:

- $i \in E_{[i]}$ and
- $f_{i}(J^{V}) = J_{a}^{U}$.

First we define a process, that collects the different set of indices $\{i_{V} \in [M]^*\}_{V \in [N]}$, such that the elements of the set makes an interval of all type. Start the process with the zeroth level $N$-tuple:

$$(3.25) \quad \mathcal{T}^{0}(\emptyset) := \{(i^{0,0}, \ldots, i^{0,N-1})\}.$$ 

and for $c_{n} \in [L]^{n}$ the level-$n$ $N$-tuple of $c_{n}$ is:

$$(3.26) \quad \mathcal{T}^{n}(c_{n}) = \left\{ \left( i_{0}^{n,0}, \ldots, i_{0}^{n,N-1} \right), \ldots, \left( i_{j}^{n,0}, \ldots, i_{j}^{n,N-1} \right) \right\},$$

in a way that the following three conditions hold:

- $i_{n,V}^{n, V} \in [M]^{n+k}$,
- $f_{i_{n,V}}(J^{V}) = J_{c_{n}}^{\tilde{U}}$, meaning that $i_{n,V}$ makes $J_{c_{n}}^{\tilde{U}}$ be of type $V$.
- For all $\ell_{1}, \ell_{2} \in [j]$ and for all $V \in [N]$: $i_{\ell_{1}}^{n,V} \neq i_{\ell_{2}}^{n,V}$, meaning that all elements appear only once.

Observe that the distribution of $\# \mathcal{T}^{n}(a_{n}) | H^{0}$ is the same as the distribution of $\min_{U \in [N]} \# S^{U}(a_{n})$. Now given $H^{0}$ consider $\# \mathcal{T}^{n}(a_{n})$. By condition (1) of the Theorem, we have $\gamma > 1$ (see (3.21)). Choose $1 < \rho < \gamma$, and define the events

$$(3.27) \quad H_{n} := \{ \forall a_{n} \in [L]^{n} : \# \mathcal{T}^{n}(a_{n}) > \rho^{n} \}.$$ 

Recall that $1 < \gamma = \min_{V \in [N]} \min_{a \in [L]} \mathbb{E}(\# S^{V}(a))$, by Lemma 3.4 and (3.20). As usual let $\overline{A}$ denote the complement of the event $A$.

**Lemma 3.6.** There exists a $0 < \delta < 1$ such that for all $n \in \mathbb{N}$

$$(3.28) \quad \mathbb{P} \left( \overline{H}_{n+1} | H_{n}, H_{0} \right) \leq L^{n+1} \cdot N \cdot \delta^{\rho^{n}}.$$ 

**Fact 3.7** (Azuma-Hoeffding inequality). Let $Z_{0}^{U}(k), \ldots, Z_{\ell}^{U}(k)$ be i.i.d random variables distributed according to $\# S^{U}(k)$, then

$$(3.29) \quad \mathbb{P} \left( Z_{0}^{U}(k) + \cdots + Z_{\ell}^{U}(k) \leq \ell \cdot \rho \right) \leq \delta^{\ell}, \text{ for some } 0 < \delta < 1.$$
Proof of Lemma 3.6.

(3.30) \( \mathbb{P}(H_{n+1}|H_n, H_0) \)
\[
\leq \sum_{a_n \in [L^n]} \sum_{a \in [L]} \mathbb{P}(\#\Xi^n(a_n) < \rho^{n+1} | \forall c_n \in [L] : \#\Xi^n(c_n) \geq \rho^n)
\]
For a fixed \( a_n \in [L]^n, a \in [L] \)

(3.31) \( \mathbb{P}(\#\Xi^n(a_n) < \rho^{n+1} | \forall c_n \in [L] : \#\Xi^n(c_n) \geq \rho^n) \)
\[
\leq \sum_{U \in [N]} \mathbb{P}(\#S^U(a_n) < \rho^{n+1} | \min_{V \in [N]} \#S^V(a_n) \geq \rho^n).
\]
For any fixed \( U \) we can use Fact 3.7 to upper-bound the last probability. This is because conditioned on we have at least \( \rho^n \) level \( n \) N-lets, \( \#S^U(a_n) \) is at least the sum of at least \( \rho^n \) independent random variables, distributed according to \( \#S^U(a) \). The reason that the random variables in this sum are independent is as follows: By the construction, given \( E_n \), the events that we retain or discard different cylinders on level \( n+1 \) are independent, hence the random variables in the sum are also indeed independent.

(3.32) \( \mathbb{P}(\#S^U(a_n) < \rho^{n+1} | \min_{V \in [N]} \#S^V(a_n) \geq \rho^n) \)
\[
\leq \delta(U, a)^{\rho^n}
\]
for some \( 0 < \delta(U, a) < 1 \), for any particular \( U \in [N], a_n \in [L]^n \) and \( a \in [L] \). Choose \( \delta := \max_{U \in [N]} \max_{a \in [L]} \delta(U, a) \), in this way we get that:

(3.33) \( \mathbb{P}(H_{n+1} | H_n) \leq \sum_{a_n \in [L^n]} \sum_{a \in [L]} \sum_{U \in [N]} \delta^{\rho^n} = L^{n+1} \cdot N \cdot \delta^{\rho^n} \)
\[
\square
\]

Lemma 3.8. For every \( n \in \mathbb{N} \):

(3.34) \( \mathbb{P}(\forall V \in [N], \forall a_n \in [L]^n : \#S^V(a_n) \geq \rho^n) > 0 \)

Proof. This is an adaptation of Dekking and Simon ([6]) Lemma 1. For \( a_n \in [L]^n \):

(3.35) \( e^T A_{a_n} \geq \rho^n e^T A_{a_n} = [CS_{a_n,0}, \ldots, CS_{a_n,N-1}] \geq [\rho^n, \ldots, \rho^n]. \)

This means that in the deterministic setup we have enough indices for the event to happen for any \( a_n \). Hence it follows that \( \#S^V(a_n) \) happens with positive probability for each \( a_n \) and \( V \), and as the events are not exclusive the intersection of them also has positive probability. \( \square \)
It is easy to see, that if we can prove that

\[(3.36) \quad \mathbb{P}(\forall n, \forall a_n : \#T^n(a_n) > 0) > 0\]

then we prove that \(\Lambda_F(p)\) contains an interval with positive probability. This is because (3.36) means that with positive probability for any \(n\) we retain something in every sub-interval \(J_{\tilde{k}a_n}^{\tilde{U}}\) (for all \(a_n \in [L]^n\)) of \(J_{\tilde{k}}^{\tilde{U}}\). Then (3.36) holds, since

\[(3.37) \quad \mathbb{P}(H_n, n > r) = \mathbb{P}(H_0) \cdot \mathbb{P}(H_r) \prod_{n=r}^{\infty} \mathbb{P}(H_{n+1} | H_n) \geq p_0 \cdot \mathbb{P}(H_r) \prod_{n=r}^{\infty} (1 - L^{n+1} \cdot N \cdot \delta^{n}) .\]

\(p_0\) is positive by Lemma 3.5, \(\mathbb{P}(H_r)\) is positive by Lemma 3.36, and we can choose \(r\) such that the last expression is positive. Hence

\[(3.38) \quad \mathbb{P}(\Lambda_F(p) \text{ contains an interval}) > 0.\]

**Lemma 3.9.** Let \(\theta\) be a possible property of \(\Lambda_F(p)\). Let \(\mathcal{A}\) denote the event that \(\theta\) holds for \(\Lambda_F(p)\). Assume that \(\mathcal{A}\)

1. happens almost surely if the process dies out,
2. happens if and only if \(\theta\) holds for every intersection of the set with level \(n\) cylinders, namely for every \(n\) and \(i_1, \ldots, i_n\): \(\theta\) holds for \(\Lambda_F(p) \cap I_{i_1, \ldots, i_n}\) and
3. is not a sure event.

Then conditioned on non-extinction \(\mathcal{A}\) almost surely does not happen.

**Proof.** The proof is based on a standard argument (similar to the one in [10, page 471]) using statistical self-similarity. Let \(\mathcal{C} := \{\#E_\infty > 0\}\) denote the event that the process does not die out. By the third assumption of the lemma

\[(3.39) \quad \mathbb{P}(\mathcal{A}) = \varepsilon < 1.\]

From the theory of branching processes we know that conditioned on non-extinction the number of retained level \(n\) cylinders tends to infinity almost surely, i.e. \(\mathbb{P}(\#E_n < \tilde{M}^n | \mathcal{C}) \to 0\) as \(n \to \infty\) for all \(\tilde{M}\). Now, by the
assumption of the Lemma:
\[
\mathbb{P} (\mathcal{A} \mid \mathcal{C}) \\
= \mathbb{P} (\mathcal{A} \mid \#\mathcal{E}_n \geq \tilde{M}) \cdot \mathbb{P} \left( \#\mathcal{E}_n \geq \tilde{M} \mid \mathcal{C} \right) \\
+ \mathbb{P} (\mathcal{A} \mid \#\mathcal{E}_n < \tilde{M}) \cdot \mathbb{P} \left( \#\mathcal{E}_n < \tilde{M} \mid \mathcal{C} \right) \\
\leq \varepsilon^{\tilde{M}} + \mathbb{P} \left( \#\mathcal{E}_n < \tilde{M} \mid \mathcal{C} \right).
\]

(3.40)

The second inequality follows from the second condition of the lemma
\[
\mathcal{A} = \{ \forall n, \forall i_1, \ldots, i_n \in \mathcal{E}_n : \theta \text{ holds for } \Lambda_F(p) \cap I_{i_1, \ldots, i_n} \}
\]
and statistical self-similarity. Namely from the fact that conditioned on \( \mathcal{E}_n \) in all the retained level \( n \)-cylinders we have independent rescaled copies of \( \Lambda_F(p) \). First let \( n \to \infty \), then \( \tilde{M} \to \infty \) gives that
\[
\mathbb{P} (\mathcal{A} \mid \mathcal{C}) = 0.
\]

\[
\square
\]

Proof of Theorem 1.9. Follows from (3.38) and Lemma 3.9, using that does not containing an interval is property that satisfies the assumptions of Lemma 3.38.

\[
\square
\]

3.3. Proof of Theorem 1.10.

Proof of Theorem 1.10. Parts of the following proof resembles [6, Proof of Thorem 1 (b)]. Let \( a \in [L] \) such that the spectral radius of \( p \cdot A_a \) is smaller than 1. Solely for this section let \( \tilde{a} \) denote the \( n \)-long vector consisting only of \( a \) i.e.
\[
\tilde{a}^n := (a, \ldots, a).
\]

It is a well known fact that if the spectral radius of an \( N \times N \) matrix \( B \) is smaller than 1, then
\[
\lim_{n \to \infty} \|B^n\| = 0.
\]

(3.43)

This implies by the assumptions of the theorem, that \( \lim_{n \to \infty} \|(p \cdot A_a)^n\| = \lim_{n \to \infty} \|p^n \cdot A_{\tilde{a}^n}\| = 0 \). By the sub-multiplicativity of the \( \| \cdot \| \) matrix norm defined in (3.6), it follows that for any \( \xi_k = (c_1, \ldots, c_k) \in [L]^j \):
\[
\lim_{n \to \infty} \|p^{n+k} A_{\xi_k \tilde{a}^n}\| \leq \lim_{n \to \infty} \|p^k A_{\xi_k}\| \cdot \|p^n A_{\tilde{a}^n}\| = \|p^k A_{\xi_k}\| \lim_{n \to \infty} \|p^n A_{\tilde{a}^n}\| = 0.
\]

(3.44)
Let \( c_k \in [L]^k \) be given and \( Z_n \) denote the number of level \( k+n \) cylinders intersecting \( \bigcup_{j \in [N]} J_{\Xi \xi \xi}^j \). From Lemma 3.4 we know that
\[
E(Z_n) = \| p^{k+n} A_{\Xi \xi \xi} \|,
\]
and hence from (3.44), it follows that
\[
\lim_{n \to \infty} E(Z_n) = 0,
\]
thus by Markov’s inequality:
\[
P(Z_n \geq 1) \leq E(Z_n) \to 0 \text{ as } n \to \infty.
\]
in this way the points
\[
\bigcup_{j \in [N]} \bigcap_{n \to \infty} J_{\Xi \xi \xi}^j
\]
are not contained in \( \Lambda_{\Xi}(p) \) with probability one. By varying \( c_k \) we get a countable dense set which is not contained in \( \Lambda_{\Xi}(p) \) with probability one, hence it cannot contain an interval. \( \square \)

3.4. **Proof of Theorem 1.11.**

**Proposition 3.10.** Under the conditions of Theorem 1.11 there exists a set \( K \) of positive \( \text{Leb}_1 \) measure such that
\[
\mathbb{P}(x \in \Lambda_{\Xi}(p)) > 0 \quad \text{for } \text{Leb}_1-\text{a.e. } x \in K.
\]

First we prove Theorem 1.11 using Proposition 3.10 and Lemma 3.9, then we prove Proposition 3.10.

**Proposition 3.11.** \( \mathbb{P}(\text{Leb}_1(\Lambda_{\Xi}(p)) > 0) > 0. \)

**Proof of Proposition 3.11 assuming Proposition 3.10.** Since
\( \mathbb{P}(\text{Leb}_1(\Lambda_{\Xi}(p)) > 0) > 0 \) if and only if \( E(\text{Leb}_1(\Lambda_{\Xi}(p))) > 0 \) we will prove the second.
\[
E(\text{Leb}_1(\Lambda_{\Xi}(p))) = \int_{\Omega} \text{Leb}_1(\Lambda_{\Xi}(p, \omega)) \, d\mathbb{P}(\omega)
\]
\[
\geq \int_{\Omega} \int_{K} \mathbb{1}\{x \in \Lambda_{\Xi}(p, \omega)\} \, d\text{Leb}_1(x) \, d\mathbb{P}(\omega)
\]
\[
= \int_{K} \int_{\Omega} \mathbb{1}\{x \in \Lambda_{\Xi}(p, \omega)\} \, d\mathbb{P}(\omega) \, d\text{Leb}_1(x)
\]
\[
\geq \int_{K} \mathbb{P}(x \in \Lambda_{\Xi}(p)) \, d\text{Leb}_1(x) > 0,
\]
where the last inequality follows from Proposition 3.10. \( \square \)
Proof of Theorem 1.11 assuming Proposition 3.10. By a similar argument as in the proof of Theorem 1.9 using Proposition 3.11 and Lemma 3.9. □

The proof of Theorem 1.11 assuming Proposition 3.10 is very similar to the one presented in [13, Proof of Theorem 2., page 140] hence we will not repeat this proof. In what follows we prove Proposition 3.10, using the same –branching process method – as in [13], but with a modified process.

Proof of Proposition 3.10. We will use the first part of the proof of Theorem 1.9, namely the part until (3.24). We choose \( \tilde{U} \) and \( \tilde{b}_k \in [L]^k \) in a way that lemma 3.5 holds for \( U = \tilde{U} \) and \( a = \tilde{b}_k \). Let

\[
K := J^* \tilde{U} \tilde{b}_k,
\]

Let \( U \sim \text{Uniform}(K) \) and \( \mathfrak{P} \) and \( \mathfrak{E} \) denote the corresponding distribution and expectation respectively. In what follows we will prove that

\[
\mathfrak{P} (\mathfrak{P} (U \in K) > 0) = 1,
\]

using the theory of branching processes in random environment. We begin with defining the process. Namely, for a given \( a = (a_1, \ldots, a_n, \ldots) \in \Sigma^{(L)} \) let

\[
\tilde{x}^0 (a) := (i^{0,0}, \ldots, i^{0,N-1}),
\]

where \( i^{0,0}, \ldots, i^{0,N-1} \) was defined in (3.24). Lemma 3.5 guarantees that such \( i^{0,0}, \ldots, i^{0,N-1} \) exists since the second condition of Theorem 1.11 is stronger than the second assumption of Theorem 1.9. Recursively if we have

\[
\tilde{x}^{t-1} (a) = \left\{ \left( i_1^{t-1,0}, \ldots, i_1^{t-1,N-1} \right), \ldots, \left( i_s^{t-1,0}, \ldots, i_s^{t-1,N-1} \right) \right\},
\]

then \( (i^{t,0}, \ldots, i^{t,N-1}) \) with elements \( i^{t,V} \in [M]^t \) is in \( \tilde{x}^t (a) \) if and only if, for some \( \Delta \) defined later the following holds:

- \( |i^{t,V}| = \Delta \cdot t + k \)
- \( i^{t,V}|_{\Delta(t-1)+k} \in \tilde{x}^{t-1} (a) \)
- \( f_{i^{t,V}} (J^V) = J^* \tilde{U} \tilde{b}_k a_{t+k} \)
- If \( i^{t,V} \) is in an element of \( \tilde{x}^t \), then it is not contained in another element of \( \tilde{x}^t \).

The elements of \( \tilde{x}^n (\bar{\theta}) \) are denoted by \( (i^{n,0}, \ldots, i^{n,N-1}) \), are called level-\( n \) \( N \)-tuples. Now we define the environment:

\[
\bar{\theta} = (\theta_0, \ldots, \theta_n, \ldots), \text{ where } \theta_k = (a_k \Delta_{+1}, \ldots, a_{(k+1)\Delta}),
\]
and the branching process in random environment is the following:

\[ Z_0 (\bar{\theta}) := \# \tilde{T}^0 (a) = 1, \]

\[ Z_n (\bar{\theta}) := \# \tilde{T}^n (a). \]

The above defined \( Z_n (\bar{\theta}) \) process is indeed a branching process in random (i.i.d. hence stationary and ergodic) environment. This is because

\[ Z_{n+1} (\bar{\theta}) = \sum_{i=1}^{Z_n (\bar{\theta})} X_{n,i} (\bar{\theta}), \]

where \( X_{n,i} (\bar{\theta}) \) is the number of level-\( n+1 \) \( N \)-tuples coming from the \( i \)-th level-\( n \) \( N \)-tuple in \( \tilde{T}^n (a) \). The random variables \( \{ X_{n,i} (\bar{\theta}) \}_{i=1}^{Z_n (\bar{\theta})} \) are independent, because what happens in different retained cylinders are independent of each other.

If \( Z_n (\bar{\theta}) \) does not die out for a given \( \bar{\theta} = (a_1, \ldots, a_{k\Delta+1}, \ldots) \), then conditioned on \( H^0 \) the point \( x \) which has \( L \)-adic expansion \( \tilde{b}_k, a_1, \ldots, a_{k\Delta+1}, \ldots \) shifted with the left endpoint of the interval \( J^{\tilde{U}} = [\tilde{u}_1, \tilde{u}_2] \) i.e. \( x = \tilde{u}_1 + \sum_{j=1}^{k} \tilde{b}_k L^{j-1} + \sum_{j=1}^{\infty} a_j L^{j-1+k} \) is contained in \( \Lambda_F (p) \). This is because if the process does not die out then on every level the cylinder containing \( x \) is retained, since it is of a retained type.

Denote the probability that the interval \( J^V_a \) is of every type with \( q (a, V) \):

\[ q (k, V) := \mathbb{P} \left( \forall U \in [N] \exists i \in \mathcal{E}_1 : f_i (J^U) = J^V_a \right), \quad \text{and} \]

\[ q := \min_{a \in [L]} \max_{V \in [N]} q (a, V). \]

To continue the proof of Proposition 3.10 we need the following fact

**Fact 3.12.** \( q > 0. \)

*The proof of the Fact.* Every matrix \( A_a \) has a positive row \( A_{a,V} \) by the second condition in the Theorem 1.11.

\[ \mathbb{P} \left( \{ \forall U : \# S^{V,U} (a) > 0 \} \right) = \mathbb{P} \left( \prod_{U \in [N]} \# S^{V,U} (a) > 0 \right) > 0
\]

\[ \iff \mathbb{E} \left( \prod_{U \in [N]} \# S^{V,U} (a) \right) > 0, \]

which is by independence and Lemma 3.4 equivalent to

\[ \prod_{U \in [N]} \mathbb{E} (\# S^{V,U} (a)) = \prod_{U \in [N]} p \cdot A_a (V, U) > 0. \]
The last inequality holds by the second condition Theorem 1.11. □

Let \( V \in [N] \) be arbitrary such that \( V = \max_{W \in [N]} q(a, W) \). Then we define
\[
U(a) := V.
\]

Using the Theory of branching processes in random environment (more precisely [2, Theorem 3]), under the following two conditions \( \mathbb{P} \)-almost surely the process \( Z_n(\theta) \) \( n \geq 1 \) does not die out with positive probability.

**C1:** There exists a \( c > 0 \) such that for all \( \theta : \mathbb{P}(Z_1(\theta) > 0) > c \);

**C2:**
\[
\frac{1}{L^\Delta} \sum_{(a_1, \ldots, a_\Delta) \in [L]^\Delta} \log \left( \mathbb{E}(Z_1(\theta) | \theta_0 = (a_0, \ldots, a_\Delta)) \right) > 0,
\]

where \( \Delta \) is natural number which is to be defined later, but was introduced as the length of the elements of the environmental random variable (see (3.52)).

To see that Condition **C1** holds, note that by the definition of \( q \), for some \( \varepsilon \in (0, 1) \) we have
\[
\mathbb{P}(Z_1(\theta) > 0) > p_0 \cdot q^\Delta (1 - \varepsilon).
\]

The proof of the second condition **C2** is a little bit trickier. Recall, that
\[
g_j = \left( \prod_{a \in [L]} CS_{a,j} \right)^{1/L}
\]
and denote
\[
\hat{\Gamma} = \min_{j \in [N]} g_j.
\]

By the first assumption of Theorem 1.11 \( p \cdot \hat{\Gamma} > 1 \).

**Lemma 3.13.** For any \( n \in \mathbb{N} \) and \( V \in [N] \):
\[
\mathbb{E} \left[ \log \left( \mathbb{E}(\# S^V(a_1, \ldots, a_n)) \right) \right] \geq n \cdot \log \left( \hat{\Gamma} \cdot p \right).
\]

**Proof.** The proof is by mathematical induction, namely, for \( n = 1 \):
\[
\mathbb{E} \left[ \log \left( \mathbb{E}(\# S^V(a)) \right) \right] = \frac{1}{L} \sum_{a \in [L]} \log \left( \mathbb{E}(\# S^V(a)) \right)
\]
\[
= \frac{1}{L} \log \left( \prod_{i \in [L]} p \cdot CS_{i,V} \right) = \log (p \cdot g_V) \geq \log (p \cdot \hat{\Gamma}).
\]

Assume that \( \mathbb{E} \left[ \log \left( \mathbb{E}(\# S^V(a_1, \ldots, a_n)) \right) \right] \geq n \cdot \log \left( \hat{\Gamma} \cdot p \right) \). From
\[
\mathbb{E}(\# S^{U,V}(a)) = p \cdot A_a(U, V),
\]
it follows, that

\[(3.62)\]
\[
E \left( \#S^V (a_1, \ldots, a_n, a_{n+1}) \right) = \sum_{U \in [N]} E \left( \#S^{U,V} (a_{n+1}) \right) \cdot E \left( \#S^U (a_1, \ldots, a_n) \right)
\]
\[
= \sum_{U \in [N]} \left[ \frac{A_{a_{n+1}}(U, V)}{CS_{a_{n+1}, V}} \right] \cdot \left[ p \cdot CS_{a_{n+1}, V} E \left( \#S^U (a_1, \ldots, a_n) \right) \right]
\]

by the concavity of the log function

\[(3.63)\]
\[
E \left( \#S^V (a_1, \ldots, a_n, a_{n+1}) \right) \geq \log \left( p \cdot CS_{a_{n+1}, V} \right)
\]
\[
+ \sum_{U \in [N]} \frac{A_{a_{n+1}}(U, V)}{CS_{a_{n+1}, V}} \log \left[ E \left( \#S^U (a_1, \ldots, a_n) \right) \right]
\]

Hence by independence and the linearity of the expectation, we get that

\[(3.64)\]
\[
E \left[ \log \left( E \left( \#S^V (a_1, \ldots, a_n, a_{n+1}) \right) \right) \right] = E \left[ \log \left( p \cdot CS_{a_{n+1}, V} \right) \right]
\]
\[
+ \sum_{U \in [N]} E \left[ \frac{A_{a_{n+1}}(U, V)}{CS_{a_{n+1}, V}} \right] E \left[ \log \left( E \left( \#S^U (a_1, \ldots, a_n) \right) \right) \right].
\]

The first part of the sum is

\[(3.65)\]
\[
E \left( \log \left( p \cdot CS_{a_{n+1}, V} \right) \right) = \log \left( g_V \cdot p \right) \geq \log \left( \hat{\Gamma} \cdot p \right),
\]

and for the second, by the induction hypothesis

\[
E \left[ \log \left( E \left( \#S^U (a_1, \ldots, a_n) \right) \right) \right] \geq n \cdot \log \left( \hat{\Gamma} \cdot p \right).
\]

Hence

\[(3.66)\]
\[
\sum_{U \in [N]} E \left[ \frac{A_{a_{n+1}}(U, V)}{CS_{a_{n+1}, V}} \right] E \left[ \log \left( E \left( \#S^U (a_1, \ldots, a_n) \right) \right) \right]
\]
\[
= n \cdot \log \left( \hat{\Gamma} \cdot p \right) \sum_{U \in [N]} E \left[ \frac{A_{a_{n+1}}(U, V)}{CS_{a_{n+1}, V}} \right] = n \cdot \log \left( \hat{\Gamma} \cdot p \right)
\]

combining the two above gives the required inequality. \(\square\)

It follows from

\[(3.67)\]
\[
E \left( Z_1 (\tilde{\theta}) \mid \theta_0 = (a_1, \ldots, a_{\Delta}) \right) \geq E \left( \#S^{U(a_\Delta)} (a_1, \ldots, a_{\Delta}) \right) q(a_{\Delta}, U(a_{\Delta}))
\]
\[
\geq q \cdot E \left( \#S^{U(a_\Delta)} (a_1, \ldots, a_{\Delta}) \right),
\]
that

\begin{equation}
\mathcal{E} \left[ \log \left( \mathbb{E} \left( Z_1 (\vartheta) | \theta_0 = (a_1, \ldots, a_\Delta) \right) \right) \right] \geq \mathcal{E} \left( \log \left( \mathbb{E} \left( \# S^V \left( (a_1, \ldots, a_\Delta) \right) \right) \right) + \log (q) \right) \geq \Delta \cdot \log \left( \hat{\Gamma} \cdot p \right) + \log (q).
\end{equation}

Since \( \log \left( \hat{\Gamma} \cdot p \right) > 0 \) we can choose \( \Delta \), such that

\begin{equation}
(3.69) \quad \Delta \cdot \log \left( \hat{\Gamma} \cdot p \right) + \log (q) > 0.
\end{equation}

\[ \square \]

3.5. Proof of Theorem 1.12. Throughout this section we always assume that the deterministic attractor, denoted by \( \Lambda_F \), has positive Lebesgue measure, \( \mathcal{L} \mathbb{e} b_1 (\Lambda_F) > 0 \). Whenever it has zero \( \mathcal{L} \mathbb{e} b_1 \)-measure the random attractor also has 0 measure hence the theorem trivially holds. In this section we consider the deterministic system and first state a theorem (Theorem 3.14) about it using that we prove Theorem 1.11 and then using similar methods as in [3] and [15] we prove Theorem 3.14. On the space \( \Sigma^{(L)} \) we define the left shift operator, \( \sigma : \Sigma^{(L)} \to \Sigma^{(L)} \):

\begin{equation}
(3.70) \quad \sigma (a) := a_2 \ldots a_n \ldots,
\end{equation}

for \( a = a_1, \ldots, a_n, \ldots \in \Sigma^{(L)} \). Also we define the uniform measure

\begin{equation}
(3.71) \quad \tilde{\mathbb{L}} := \left( \frac{1}{L}, \ldots, \frac{1}{L} \right)^N,
\end{equation}

on \( \Sigma^{(L)} \).

**Theorem 3.14.** There exist a \( w < \log \frac{M}{L} \) such that

\begin{equation}
(3.72) \quad \lim_{n \to \infty} \frac{1}{n} \log \| A_{a|n} \| = w, \quad \text{for } \tilde{\mathbb{L}} \text{ almost every } a \in \Sigma^{(L)}.
\end{equation}

The proof of this theorem is presented at end of this section.

We denote by \( t_j \) the translation that translates \([0, L] \) into \( J^j \), \( j \in [N] \). Namely, if \( J^j = [k \cdot L, (k + 1) \cdot L] \), then \( t_j (x) = x + k \cdot L \). This way for \( a \in [L]^n, n \geq 0 \) \( t_j^{-1} (J^j_a) = J^0_a \). Also let \( \Xi_j : \Sigma^{(L)} \to J^j \),

\begin{equation}
(3.73) \quad \Xi_j (a) := t_j \left( \sum_{n=1}^{\infty} a_t L^{1-n} \right) \quad \text{and} \quad \Xi (a) := \Xi_0 (a).
\end{equation}

The composition of the natural projection from the symbolic space to the interval \([0, L] \) and of the translation of \([0, L] \) into the \( j \)-th interval, \( J^j \).
Lemma 3.15. If \( x = \Xi_\ell(a) \) for some \( \ell \in [N] \) and \( a \in \Sigma^{(L)} \), then

\[
\# \{(j_1, \ldots, j_n) \in [M]^n : x \in I_{j_1, \ldots, j_n}\} \leq \|A_{a|n}\|.
\]

The proof of Lemma 3.15 is very similar to the proof of Lemma 3.21, hence we will only present the proof of the later one.

Proof of Theorem 1.12 using Theorem 3.14 and Lemma 3.15.

It follows from Theorem 3.14 and Egorov’s theorem, that for any fixed \( \varepsilon \) there exists a measurable subset \( \tilde{F}_\varepsilon \) such that

- \( \mathcal{L}(\tilde{F}_\varepsilon^c) < \varepsilon \)
- for \( a \in \tilde{F}_\varepsilon \) the convergence \( \lim_{n \to \infty} \frac{1}{n} \log \|A_{a|n}\| = w \) is uniform.

Let

\[
F_\varepsilon := \left( \bigcup_{\ell \in [N]} t_\ell(\tilde{F}_\varepsilon) \right) \cap \Lambda_\mathcal{F}.
\]

It is easy to see, that \( \text{Leb}_1(\Lambda_\mathcal{F} \setminus F_\varepsilon) < N \cdot \varepsilon \). Choose \( v \in \left( w, \log \frac{M}{L} \right) \), and \( \delta > 0 \) such that \( v = \log \left( \frac{M-\delta}{L} \right) \). Because of the uniform convergence and of the fact that \( v \) is strictly greater than \( w \), we can choose \( N' \) such that for any \( a \in \tilde{F}_\varepsilon \) and \( n > N' \):

\[
\frac{1}{n} \log \|A_{a|n}\| < v.
\]

If \( x \in F_\varepsilon \) and \( x = \Xi_\ell(a) \) for some \( \ell \in [N] \) and \( a \in \Sigma^{(L)} \) as in Lemma 3.15, then

\[
\frac{1}{n} \log \# \{(j_1, \ldots, j_n) \in [M]^n : x \in I_{j_1, \ldots, j_n}\} \leq \frac{1}{n} \log \|A_{a|n}\| < v.
\]

Hence

\[
\# \{(j_1, \ldots, j_n) \in [M]^n : x \in I_{j_1, \ldots, j_n}\} < \left( \frac{M-\delta}{L} \right)^n.
\]

As in 3.13, let

\[
\mathcal{E}_n = \mathcal{E}_n(\omega) = \{(j_1, \ldots, j_n) \in [M]^n : I_{j_1, \ldots, j_n} \text{ is selected}\},
\]

\[
\mathcal{E}_n(\omega, x) = \{(j_1, \ldots, j_n) \in \mathcal{E}_n : x \in I_{j_1, \ldots, j_n}\} \quad \text{and}
\]

\[
\mathcal{G}_n(U) = \{(j_1, \ldots, j_n) \in [M]^n : U \cap [j_1, \ldots, j_n] \neq \emptyset\}.
\]

Fix now a \( p < \frac{L}{M-\delta} \). With this choice of \( p \)

\[
\frac{M-\delta}{L} \cdot p < 1.
\]
Recall that $\Pi^{(M)}$ denotes the natural projection from the symbolic space $\Sigma^{(M)} = [M]^\mathbb{N}$ to $\mathbb{R}$.

\begin{equation}
(3.79) \quad \mathbb{E}(\#E_n(x)) = \sum_{(j_1,\ldots,j_n) \in G_n((\Pi^{(M)})^{-1}(x))} \mathbb{P}(I_{j_1,\ldots,j_n} \text{ is selected}) = \#G_n((\Pi^{(M)})^{-1}(x)) \cdot p^n < \left( \frac{M - \delta}{L} \cdot p \right)^n =: b^n.
\end{equation}

Let
\begin{equation}
(3.80) \quad u := u(\varepsilon) := \mathcal{L}eb_1(F_\varepsilon \cap \Lambda_\mathcal{F}).
\end{equation}

Recall, that our assumption was that $\mathcal{L}eb_1(\Lambda_\mathcal{F}) > 0$. Let
\[ \hat{\mu} := \frac{\mathcal{L}eb_1|_{F_\varepsilon \cap \Lambda_\mathcal{F} \times \mathbb{P}}}{u}, \]
and consider the finite measure-space: $(F_\varepsilon \times \Omega, \mathcal{B} \times \mathcal{A}, \hat{\mu})$ ($\mathcal{B}$ denotes the Borel sigma-algebra and recall: $(\Omega, \mathcal{A})$ was defined in 3.9) and let $\Lambda_\mathcal{F}(p, \omega)$ denote the random Cantor set corresponding to the realization $\omega$. Then by the previous calculation,
\begin{equation}
(3.81) \quad \int_{x \in F_\varepsilon} \int_{\omega \in \Omega} \#E_n(\omega, x) d\hat{\mu}(\omega, x) \leq b^n.
\end{equation}

**Lemma 3.16.** For $\mathbb{P}$ almost every $\omega \in \Omega$ we have
\[ \mathcal{L}eb_1(\Lambda_\mathcal{F}(p, \omega) \cap F_\varepsilon) = 0. \]

**Proof.** Assume that there exists an $\hat{\varepsilon} > 0$ such that
\begin{equation}
(3.82) \quad \mathbb{P}(\mathcal{L}eb_1(\Lambda_\mathcal{F}(p, \omega) \cap F_\varepsilon) > \hat{\varepsilon)) > \hat{\varepsilon}.
\end{equation}

The $n$-th approximation of the random attractor is denoted by
\[ \Lambda_{n,\mathcal{F}}(p, \omega) := \bigcup_{(j_1,\ldots,j_n) \in E_n(\omega)} I_{j_1,\ldots,j_n}. \]

Choose $n > N'$ (such that (3.75) holds). By the assumption (3.82) $\mathbb{P}(\omega \in \Omega : \mathcal{L}eb_1(\Lambda_{n,\mathcal{F}}(p, \omega) \cap F_\varepsilon) > \hat{\varepsilon)) > \hat{\varepsilon}$ also holds. Let
\begin{equation}
(3.83) \quad H_n := \{ \omega \in \Omega : \mathcal{L}eb_1(\Lambda_{n,\mathcal{F}}(p, \omega) \cap F_\varepsilon) > \hat{\varepsilon) \}.
\end{equation}

Then $\mathbb{P}(H_n) > \hat{\varepsilon}$. For an $\omega \in H_n$, by the definition of $\Lambda_{n,\mathcal{F}}(p, \omega)$
\[ F_\varepsilon \cap \Lambda_{n,\mathcal{F}}(p, \omega) = \bigcup_{(j_1,\ldots,j_n) \in E_n(\omega) \cap \Lambda_{n,\mathcal{F}}(p, \omega)} I_{j_1,\ldots,j_n} \cap F_\varepsilon. \]
It follows from the assumption (3.82) that

\[(3.84)\]
\[\hat{\varepsilon} < \mathcal{L}eb_1(\Lambda_\mathcal{F}(p, \omega) \cap F_\varepsilon) \leq \int_{F_\varepsilon} \sum_{(j_1, \ldots, j_n) \in E_n(\omega)} \mathbb{I}_{I_{j_1, \ldots, j_n}(x)} \, dx = \int_{F_\varepsilon} \# \mathcal{E}_n(\omega, x) \, dx\]

Since \(n > N'\) by (3.79), \(\mathbb{E}(\# \mathcal{E}_n(x)) < b^n\), and since \(b < 1\), \(b^n\) can be arbitrarily small, which contradicts the above calculation, which shows that \(\mathbb{E}(\# \mathcal{E}_n(x)) > \hat{\varepsilon}\). □

**Proof of Theorem 1.12.** Let \(\tilde{\varepsilon} = \frac{1}{k}\) for \(k \in \mathbb{N}\). Then there exists a set \(F_1^k\) such that \(\mathcal{L}eb_1(\Lambda_\mathcal{F} \setminus F_1^k) < N \cdot \frac{1}{k}\), and \(\Omega_k \subset \Omega\) with \(\mathbb{P}(\Omega_k) = 1\) such that for any \(\omega \in \Omega_k\): \(\mathcal{L}eb_1(\Lambda_\mathcal{F}(p, \omega) \cap F_1^k) = 0\). It follows that \(\mathcal{L}eb_1(\Lambda_\mathcal{F}(p, \omega)) < \frac{N}{k}\) holds for any \(\omega \in \Omega_k\). Let \(\widetilde{\Omega} = \bigcap_{k=1}^{\infty} \Omega_k\) then \(\mathbb{P}(\widetilde{\Omega}) = 1\) and for any \(\omega \in \widetilde{\Omega}\) we get \(\mathcal{L}eb_1(\Lambda_\mathcal{F}(p, \omega)) = 0\). □

The rest of this section is devoted to the proof of Theorem 3.14. It is a simple consequence of a combination of a number of theorems due to Ruiz [15] and jointly to Bárány and Rams [3]. First we state these assertions and then we provide the proof of Theorem 3.14.

**Fact 3.17.** For any \(\ell \in [N], b \in [L]\) and \(a \in [L]^n\), \(n \geq 0\):

\[(3.85)\]
\[\nu(J_{ba}^\ell) = \sum_{k \in [N]} \frac{1}{M} A_b(\ell, k) \nu(J_a^k).\]

This fact can be proved by the same argument used in [15, bottom of page 354].

For an \(a \in [L]^n\) we form the vectors:

\[(3.86)\]
\[\nu(. , a) = (\nu(J_a^0), \ldots, \nu(J_a^{N-1})), \quad \nu(. , \emptyset) = (\nu(J^0), \ldots, \nu(J^{N-1})).\]

Fact 3.17 implies that \(\nu(J_a^\ell) = \frac{1}{M^n} e_\ell^T \cdot A_a \nu(. , \emptyset),\) where \(e_\ell \in \mathbb{R}^N\) is the \(\ell\)-th coordinate unit vector. Let

\[(3.87)\]
\[\eta_j := \nu_j \circ t_j, \quad \eta := \sum_{j \in [N]} \eta_j\]

where \(\nu_j = \nu_{J_j}\).

It is easy to see, that \(\eta\) is probability measure on \([0, L]\). We also define

\[(3.88)\]
\[\tilde{\eta}(a_1, \ldots, a_n) := \eta(J_a^0, \ldots, a_n).\]
Observe that

\begin{equation}
\eta(J_0^a) = \sum_{j \in [N]} \nu(J_0^k) = \frac{1}{M^n} e^T A_0 \nu(\cdot, \emptyset).
\end{equation}

Now we briefly explain the intuition behind the method presented below. Instead of studying the original attractor we consider the one which we get by intersecting the original attractor with the different intervals $J^0, \ldots, J^N$, and translate all of the resulting intersection sets to $[0, 1]$. The measure $\eta$ can be thought of as the natural measure of the modified system corresponding to the modified attractor. Also to be able to connect the matrices $A_a$ and their products to the behaviour of the system it is more convenient for us to work in the symbolic space $\Sigma(L)$ containing the $L$-adic codes of the points of $[0, L]$ (the support of our new attractor).

**Lemma 3.18.** The matrix $A = \sum_{k \in [N]} A_k$ is primitive, meaning that there exists a $K > 0$ such that $A^K(i, j) > 0$ for every $i, j \in \{1, \ldots, N\}$.

**Proof.** It is proven in [4, Section 4.4], but for the convenience of the reader we present the proof of this lemma.

Recall that $|I| = \tilde{n}L$. Let $d_\ell$ be the distance between $x_\ell$ and the nearest end-point of $J_\ell$, let $d' > 0$ be the minimum of the $d_\ell$s. We choose $K$ such that $|f_{i|}^\ell(I)| = \frac{\tilde{n}}{L^{K-1}} < d'$. Using this and the fact that $x_\ell \in f_{i|}^\ell(I)$ we obtain that $f_{i|}^\ell(I) \subset J^\ell$. It follows that for any $u \in [N]

\begin{equation}
M_{i|}^\ell(J^u) \subset J^\ell.
\end{equation}

Hence there exists an $a^u \in [L]^K$ such that $f_{i|}^\ell(J^u) = J_0^a$. It follows that $A_{i|}^\ell(\ell, u) > 0$ and thus for every $\ell$ and $u$

\begin{equation}
A^K(\ell, u) = \sum_{i_1, \ldots, i_K} A_{i_1, \ldots, i_K}(\ell, u) > 0.
\end{equation}

\[\Box\]

**Lemma 3.19.** $\tilde{\eta}$ is $\sigma$-invariant and mixing.

**Proof.** The proof is a slightly modified version of the proof in [3, Lemma 3.4]. In our case instead of using [3, Lemma 3.1] we use the previous lemma – Lemma 3.18.

\[\Box\]

**Lemma 3.20.** $\text{dim}_H \tilde{\eta} < 1$. 

\[\Box\]
Proof. The fact that \( \dim_H \nu < 1 \) holds was made explicit first in [18, Theorem 44] and the proof of this fact is also available in [4, Section 4.4.4]. On the other hand, \( \dim_H \tilde{\eta} = \dim_H \nu \) follows from the arguments in [15, page 357]. \( \square \)

Recall that we denoted \([M]^N\) by \( \Sigma^\(M\)\). Moreover, \( \Sigma^\(M\)\) is equipped with the metrics
\[
\rho(i, j) := L^{-|i \land j|}.
\]

For any \( U \subset \Sigma^\(M\)\) we have
\[
(3.89) \quad \overline{\dim}_B(U) = \lim_{n \to \infty} \frac{\log \#G_n(U)}{n \log L},
\]
where \( G_n(U) \) was defined in (3.78). Let \( \tau : \Sigma^(L) \to \Sigma^\(M\)\),
\[
(3.90) \quad \tau(a) := (\Pi^\(M\))^\(-1\) \left( \bigcup_{\ell \in [N]} \Xi_\ell(a) \right).
\]

Now we briefly explain the meaning of this function. Recall that \( \Xi_\ell \) is the composition of a projection from the symbolic space \( \Sigma^\(L\)\) to \([0, L]\) and a translation to the \( \ell \)-th interval, \( J^\ell \). Hence what we do with \( \tau \) is that we take an \( L \)-adic representation \( a \in \Sigma^L \), then consider the point \( x \) with \( L \)-adic representation \( a \) and translate it with the left-endpoints of the intervals \( J^1, \ldots, J^N \) to get the points \( x_1, \ldots, x_N \in \mathbb{R} \). Then consider the symbolic space corresponding the iterated function system, \( \Sigma^\(M\)\). The resulting set consists of those points of \( \Sigma^\(M\)\) which natural projection \( (\Pi^\(M\)) \) is one of \( x_1, \ldots, x_N \).

The relations between the different projections are indicated by the following commutative diagram.
\[
(3.91)
\]

In the following section our goal is to upper bound the upper box dimension of the set \( \tau(a) \) for a large set of \( a \in \Sigma^\(L\)\). Recall that for any \( a \in \Sigma^\(L\)\):
\[
\overline{\dim}_B(\tau(a)) = \limsup_{n \to \infty} \frac{\log \#G_n(\tau(a))}{n \log L}.
\]
In what follows we give an upper bound on \( \#G_n(\tau(a)) \).
Lemma 3.21. For any \( a = (a_1, \ldots, a_n, \ldots) \) we have \( \#G_n(\tau(a)) \leq ||A_{a_1, \ldots, a_n}|| \).

Proof. Fix \( a \in \sigma^{(M)} \). Recall that the definition of \( G_n(\tau(a)) \) is the following:

\[
G_n(\tau(a)) = \{ (j_1, \ldots, j_n) \in [M]^n : [j_1 \ldots j_n] \cap \tau(a) \neq \emptyset \}.
\]

Hence, for a \( (j_1, \ldots, j_n) \in [M]^n \) we have \( (j_1, \ldots, j_n) \in G_n(\tau(a)) \) if there exists an \( i \in \Sigma^{(M)} \) with \( i|n = j_1, \ldots, j_n \) and an \( \ell \in [N] \) such that \( \lim_{k \to \infty} f_{i_k}(0) = \Xi_{\ell}(a) \in J^\ell_{a_1, \ldots, a_n} \). It follows that \( \Xi_{\ell}(a) \in f_{j_1 \ldots j_n}(\Lambda) \subset f_{j_1 \ldots j_n}(\bigcup_{k \in [N]} J^k) \). Therefore there exists a \( k \) such that \( \Xi_{\ell}(a) \in f_{j_1 \ldots j_n}(J^k) \). Since for any interval \( J^k \) there exists \( \ell \) and \( b = (b_1, \ldots, b_n) \) such that \( f_{j_1 \ldots j_n}(J^k) = J^\ell_b \), it follows from \( J^\ell_{a_1} \cap f_{j_1 \ldots j_n}(J^k) \neq \emptyset \), that \( J^\ell_{a_1} = f_{j_1 \ldots j_n}(J^k) \). Let

\[
U_n^{k,\ell}(a) := \# \{ (j_1, \ldots, j_n) \in [M]^n : f_{j_1 \ldots j_n}(J^k) = J^\ell_{a_1} \}.
\]

It is easy to see from the definition of \( A_{a_1, \ldots, a_n} \) and the argument presented above \( U_n^{k,\ell}(a) \) and the above that

\[
\begin{align*}
\# \{ (j_1, \ldots, j_n) \in [M]^n : [j_1 \ldots j_n] \cap \tau(a) \neq \emptyset \} & \leq \# \{ (j_1, \ldots, j_n) \in [M]^n : \exists k, \exists \ell f_{j_1 \ldots j_n}(J^k) = J^\ell_{a_1, \ldots, a_n} \} \\
& \leq \# \bigcup_{k \in [N]} \bigcup_{\ell \in [N]} \{ (j_1, \ldots, j_n) \in [M]^n : f_{j_1 \ldots j_n}(J^k) = J^\ell_{a_1, \ldots, a_n} \} \\
& \leq \sum_{k \in [N]} \sum_{\ell \in [N]} U_n^{k,\ell}(a) = ||A_{a_1, \ldots, a_n}||
\end{align*}
\]

It follows from Lemma 3.21, that

\[
\dim_B(\tau(a)) \leq \limsup_{n \to \infty} \frac{\log(||A_{a_1, \ldots, a_n}||)}{n \log(L)}
\]

Let us define the pressure function \( P(t) \) as

\[
P(t) := \lim_{n \to \infty} \frac{1}{n \log(L)} \log \sum_{(a_1, \ldots, a_n) \in [L]^n} (||A_{a_1, \ldots, a_n}||)^t.
\]

Since \( ||L^n = L^n \), it is easy to see, that \( P(0) = 1 \). Also, \( \sum_{(a_1, \ldots, a_n) \in [L]^n} (||A_{a_1, \ldots, a_n}||) = ||A^n|| \) and from the meaning of the matrices \( A \) and \( A^n \), it follows that \( ||A^n|| = M^n \), hence \( P(1) = \frac{n \log M}{n \log L} = s \).

Lemma 3.22. The function \( t \to P(t) \)

- exists for \( t \in \mathbb{R} \),
- is monotone increasing, convex, continuous,
- is continuously differentiable for \( t > 0 \).
Proof. Since the matrix $A$ is primitive (see Lemma 3.18), the proof is similar to the proof of Lemma 4.3 in [3]. □

The following lemma is also a slightly modified version of lemma 4.5. in [3].

**Lemma 3.23.** There exists a unique ergodic, left-shift invariant Gibbs measure $\mu_1$ on $\Sigma^{(L)}$ such that, there exists a $C > 0$ that for any $(a_1, \ldots, a_n) \in [L]^n$ we have

1. $C^{-1} \leq \frac{\mu_1([a_1, \ldots, a_n])}{\|A_{a_1 \ldots a_n}\|^{M-n}} \leq C$;
2. $\dim_H \mu_1 = -P'(1) + s$;
3. $\lim_{n \to \infty} \frac{\log \|A_{a_1 \ldots a_n}\|}{n \log(L)}$.

From lemma 3.19 we know that $\tilde{\eta}$ is also an ergodic measure on $\Sigma^{(L)}$, hence it follows, that $\tilde{\eta} = \mu_1$.

Hence by Lemma 3.20 $\dim_H \mu_1 = \dim_H \tilde{\eta} < 1$ thus by the second part of Lemma 3.23

$P'(1) > s - 1$.

From the convexity and continuity of the pressure function it follows that

$$\inf_{t > 0} (P(t) - (s - 1)t) < 1,$$

see Figure 9 for a visual explanation.

![Figure 9. Explanation of (3.92).](image)

**Lemma 3.24.**

(3.93) $\dim_H \left\{ a \in \Sigma^{(L)} : \lim_{n \to \infty} \frac{1}{n \log L} \log \|A_{a_1 a_n}\| \geq s - 1 \right\}$

$\leq \inf_{t > 0} (P(t) - (s - 1)t)$.

Proof. The proof is the same as the proof of Lemma 4.7 in [3]. □
From Lemma 3.24 and equation (3.92) it follows that
\[
\dim_H \left\{ a \in \Sigma^{(L)} : \lim_{n \to \infty} \frac{1}{n \log L} \log \| A_{a|n} \| \geq s - 1 \right\} < 1
\]

**Proof of Theorem 3.14.** Let \( \varphi_n : \Sigma^{(L)} \to \mathbb{R} \),
\[
\varphi_n(a) = \log \| A_{a|n} \|.
\]

From the properties of the 1-norm, it is easy to see, that
\[
\| A_{a_1,\ldots,a_n,a_{n+1},\ldots,a_{n+m}} \| \leq \| A_{a_1,\ldots,a_n} \| \cdot \| A_{a_{n+1}} \ldots A_{a_{n+m}} \|,
\]
hence, \( \varphi_{n+m}(a) \leq \varphi_n(a) + \varphi_m(\sigma^n a) \), i.e. the function \( \varphi \) is subadditive. Recall, that \( \tilde{\mathcal{L}} = (\frac{1}{L}, \ldots, \frac{1}{L})^N \) denotes the uniform distribution measure on \( \Sigma^{(L)} \). \( \tilde{\mathcal{L}} \) is \( \sigma \)-invariant and ergodic. Also clearly \( \varphi_1 \) is an \( L^1(\tilde{\mathcal{L}}) \) function. Hence we can use the Subadditive ergodic theorem ([19], Theorem 10.1), which states that there exists a \( \varphi \) such that

1. \( \varphi \in L^1(\tilde{\mathcal{L}}) \),
2. \( \varphi \circ \sigma = \varphi \) almost everywhere,
3. \( \lim_{n \to \infty} \frac{1}{n} \varphi_n = \varphi \) almost everywhere,
4. \( \lim_{n \to \infty} \frac{1}{n} \int \varphi_n d\tilde{\mathcal{L}} = \int \varphi d\tilde{\mathcal{L}} \).

Since \( \tilde{\mathcal{L}} \) is ergodic measure and \( \varphi(a) = \varphi(\sigma a) \) for \( \tilde{\mathcal{L}} \) almost every \( a \in \Sigma^{(L)} \) there exists a \( w \) such that \( \varphi(a) = w \) for \( \tilde{\mathcal{L}} \) almost every \( a \in \Sigma^{(L)} \). \( \square \)

4. **Appendix**

**Figure 10.** The first three level approximation of the Sierpiński carpet.

4.1. **Random Sierpiński carpet.** The (deterministic) Sierpiński carpet is the attractor (see Figure 2.1) of the following self-similar IFS in \( \mathbb{R}^2 \):

\[
\mathcal{F} := \left\{ f_i(x) = \frac{1}{3}(x) + t_i \right\}_{i=0}^8,
\]
Figure 11. Illustration of the $\hat{\text{proj}}(1, -1)$ and the $\hat{\text{proj}}(1, 0)$ projection of the first level of the Sierpiński carpet.

where $\{t_i\}_{i=0}^8$ is an enumeration of the following set

$$\left\{0, \frac{1}{3}, \frac{2}{3}\right\}^2 \setminus \left\{\left(\frac{1}{3}, \frac{1}{3}\right)\right\}$$

We obtain the random Sierpiński carpet by applying the random construction introduced in Definition 1.1 for the deterministic IFS above. We denote the random Sierpiński carpet with parameter $p$ by $S_p$.

Let $\hat{\text{proj}}(\alpha, \beta)$ ($\alpha, \beta \in \mathbb{R}$) denote the following projection to $\mathbb{R}$:

$$\hat{\text{proj}}(\alpha, \beta)(x, y) = \alpha \cdot x + \beta \cdot y,$$

which is the orthogonal projection to the line of tangent $\beta/\alpha$ rescaled. In what follows we consider the $(1, -1)$ and the $(1, 0)$ projections of the random Sierpiński carpet as it is schematically illustrated in Figure 11. The behaviour of the projections of $S_p$ is quite well known (see for example [14], [17]). In particular, about the $\hat{\text{proj}}(1, 0)$ projection we know almost everything from [5]. However, by Theorems 1.9-1.12 we can further clarify the picture, as we explain below.

Example 4.1 ($\hat{\text{proj}}(1, 0)(S_p)$). The projected and rescaled IFS is the following (see Definition 1.8): $S_i(x) = \frac{1}{3}x + t_i$, $t_i \in \{0, 1, 2\}$, with $(n_0, n_1, n_2) = (3, 2, 3)$, and basic types (see (1.17)) $J^0 = [0, 1]$, $J^1 = [1, 2]$, $J^2 = [2, 3]$.
Hence, the corresponding matrices (see (1.18)) are,

\[
D_0 = \begin{bmatrix} 3 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 0 & 0 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 & 3 & 0 \\ 0 & 2 & 0 \\ 0 & 3 & 0 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{bmatrix}.
\]

In this case the spectral radiuses of \( p \cdot D_i, i \in \{0, 1, 2\} \) are 3\( p \), 2\( p \), 3\( p \) respectively.

**Example 4.2** (\( \widehat{\text{proj}}_{(1,-)}(S_p) \)). The projected and rescaled IFS is the following (see Definition 1.8): \( \widehat{S}_i(x) = \frac{1}{5} x + t_i, t_i \in \{0, 1, 2, 3, 4\} \), with \((n_0, n_1, n_2, n_3, n_4) = (1, 2, 2, 1, 1)\), and basic types (see (1.17)) \( \widehat{J}^0 = [0, 3], \widehat{J}^1 = [3, 6] \). Hence, the corresponding matrices (see (1.18)) are,

\[
C_0 = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}.
\]

The spectral radiuses of \( p \cdot C_i (i \in \{0, 1, 2\}) \) are 2\( p \), 3\( p \), 2\( p \) respectively.

As we mentioned already the \( \widehat{\text{proj}}_{(1,0)} \) projection of \( S_p \) is heavily studied by [5]. Their results are illustrated in Figure 12. What we can add to this is as follows:

1. by Theorem 1.10 we know that for \( p < \frac{1}{2} \) the projection \( \widehat{\text{proj}}_{(1,0)} \) does not contain any intervals almost surely.

It is clear from the results in [5] that for \( p > \frac{1}{2} \) the projection \( \widehat{\text{proj}}_{(1,0)}(S_p) \) contains an interval almost surely conditioned on non-extinction.

Moreover, it is stated in [14], that this last result holds for every direction, especially for the \((-1,1)\) direction too. About the \( \widehat{\text{proj}}_{(-1,1)} \) projection we further know from [17] that there exists a \( p' > \frac{3}{8} \), such that for \( p < p' \), \( \dim_H(S_p) > 1 \) almost surely conditioned on non-extinction but the \( \widehat{\text{proj}}_{(-1,1)} \) projection does not contain any intervals almost surely. Using Theorems 1.9-1.12 we can extend these results as follows:

1. there exists a \( p'' \) such that for \( p < p'' \), \( \dim_H(S_p) > 1 \) almost surely conditioned on non-extinction but the projection \( \widehat{\text{proj}}_{(-1,1)}(S_p) \) has not only empty interior but also zero Lebesgue measure.

2. For \( p > 0.38 \cdots = \frac{1}{18^{1/3}} \) (note that this is the same value that occurred in the other projection \( \widehat{\text{proj}}_{(1,0)} \) in Figure 12) the projection \( \widehat{\text{proj}}_{(-1,1)}(S_p) \) has positive Lebesgue measure almost surely conditioned on non extinicion.
Figure 12. Illustration of the results of Dekking and Meester [5] on the \( \hat{\text{proj}}(1,0) \) phases of the projection of the random Sierpiński carpet. It follows from our results that the content of the red rectangle can be replaced with "\( \hat{\text{proj}}_p(1,0) \) has positive Lebesgue measure a.s. conditioned on non-extinction but empty interior a.s.".

(3) For \( p < \frac{1}{2} \) the projection \( \hat{\text{proj}}_{(-1,1)}(S_p) \) does not contain an interval almost surely.

The Menger sponge is the 3-dimensional analogue of the Sierpiński carpet and the projections we studied in this chapter are analogous to those three dimensional ones which we denoted by \( \text{proj}_{(1,0,0)} \) and \( \text{proj}_{(1,1,1)} \). Nonetheless unlike in the case of the Menger sponge (where the \( p \) parameter intervals were very different in case of the two projections) here the two projections seemingly shows a very similar behavior for a given \( p \).

References

[1] K.B. Athreya and PE Ney. Branching processes. Dover Publications, 2004.
[2] Krishna B. Athreya and Samuel Karlin. On branching processes with random environments: I: Extinction probabilities. The Annals of Mathematical Statistics, 42(5):1499–1520, 1971.
[3] Balázs Bárány and Michał Rams. Dimension of slices of sierpiński-like carpets. Journal of Fractal Geometry, 1(3):273–294, 2014.
[4] Balázs Bárány, Károly Simon, and Boris Solomyak. Self-similar and self-affine sets and measures. unpublished, 2021.
[5] F. Michel Dekking and Ronald W. J. Meester. On the structure of mandelbrot’s percolation process and other random cantor sets. Journal of Statistical Physics, 58:1109–1126, 1990.
[6] Michel Dekking, Károly Simon, and Balázs Székely. The algebraic difference of two random cantor sets: The larsson family. *The Annals of Probability*, 39, 01 2009.

[7] Kenneth Falconer. *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2004.

[8] Kenneth J Falconer. Random fractals. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 100, pages 559–582. Cambridge University Press, 1986.

[9] Kenneth J Falconer and Xiong Jin. Exact dimensionality and projections of random self-similar measures and sets. *Journal of the London Mathematical Society*, 90(2):388–412, 2014.

[10] KJ Falconer and GR Grimmett. On the geometry of random cantor sets and fractal percolation. *Journal of Theoretical Probability*, 5:465 – 486, 1992.

[11] Pertti Mattila. Hausdorff dimension, orthogonal projections and intersections with planes. *Annales Academiae Scientiarum Fennicae. Mathematica*, 1:227–244, 1975.

[12] R Daniel Mauldin and Stanley C Williams. Random recursive constructions: asymptotic geometric and topological properties. *Transactions of the American Mathematical Society*, 295(1):465 – 486, 1986.

[13] Péter Móra, Károly Simon, and Boris Solomyak. The lebesgue measure of the algebraic difference of two random cantor sets. *Indagationes Mathematicae-new Series - INDAGAT MATH NEW SER*, 20, 03 2009.

[14] Michał Rams and Károly Simon. Projections of fractal percolations. *Ergodic Theory and Dynamical Systems*, 35(2):530–545, 2015.

[15] Víctor Ruiz. Dimension of homogeneous rational self-similar measures with overlaps. *Journal of mathematical analysis and applications*, 353(1):350–361, 2009.

[16] Károly Simon and Lajos Vágó. Projections of mandelbrot percolation in higher dimensions. *Springer Proceedings in Mathematics and Statistics*, 92, 07 2014.

[17] Károly Simon and Lajos Vágó. Fractal percolations. *Banach Center Publications*, 115:183–196, 2018.

[18] Lidia Torma. The dimension theory of some special families of self-similar fractals of overlapping construction satisfying the weak separation property. master’s thesis, institute of mathematics, budapest university of technology and economics, 2015. [http://math.bme.hu/~lidit/Thesi/MSC_BME_TLB.pdf](http://math.bme.hu/~lidit/Thesi/MSC_BME_TLB.pdf).

[19] P. Walters. *An Introduction to Ergodic Theory*. Graduate Texts in Mathematics. Springer New York, 2000.