DIMENSIONS OF RANDOM STATISTICALLY SELF-AFFINE SIERPINSKI SPONGES IN $\mathbb{R}^k$

JULIEN BARRAL AND DE-JUN FENG

ABSTRACT. We compute the Hausdorff dimension of any random statistically self-affine Sierpinski sponge $K \subset \mathbb{R}^k$ ($k \geq 2$) obtained by using some percolation process in $[0,1]^k$. To do so, we first exhibit a Ledrappier-Young type formula for the Hausdorff dimensions of statistically self-affine measures supported on K. This formula presents a new feature compared to its deterministic or random dynamical version. Then, we establish a variational principle expressing $\dim_H K$ as the supremum of the Hausdorff dimensions of statistically self-affine measures supported on K, and show that the supremum is uniquely attained. The value of $\dim_H K$ is also expressed in terms of the weighted pressure function of some deterministic potential. As a by-product, when $k = 2$, we give an alternative approach to the Hausdorff dimension of K, which was first obtained by Gatzouras and Lalley [27]. The value of the box counting dimension of K and its equality with $\dim_H K$ are also studied. We also obtain a variational formula for the Hausdorff dimensions of some orthogonal projections of K, and for statistically self-affine measures supported on K, we establish a dimension conservation property through these projections.

RÉSUMÉ. Nous calculons les dimensions de Hausdorff des tapis de Sierpinski aléatoires auto-affines en loi dans $\mathbb{R}^k$ ($k \geq 2$) obtenus grâce à certains processus de percolation dans $[0,1]^k$. Etant donné un tel ensemble K, en premier lieu nous exhibons une formule de type Ledrappier-Young pour les dimensions de Hausdorff des mesures statistiquement auto-affines portées par K; cette formule présente une nouvelle caractéristique comparée aux cas où K est déterministe ou associé à une dynamique aléatoire. Puis nous établissons un principe variationnel qui donne $\dim_H K$ comme le supremum des dimensions de ces mesures, et montrons que ce supremum est atteint de façon unique. La valeur de $\dim_H K$ est aussi donnée comme la pression pondérée d'un potentiel déterministe. Cette approche fournit une alternative à celle développée par Gatzouras et Lalley [27] dans le cas où $k = 2$. Nous étudions également la valeur de la dimension de boîte de K et son égalité avec $\dim_H K$. Nous obtenons aussi un principe variationnel pour les dimensions de Hausdorff de certaines projections orthogonales de K, et pour les mesures auto-similaires en loi portées par K, nous établissons une propriété de conservation de dimension au travers de ces projections.

1. INTRODUCTION

This paper deals with dimensional properties of a natural class of random statistically self-affine sets and measures in $\mathbb{R}^k$ ($k \geq 2$), namely random Sierpinski sponges and related Mandelbrot measures, as well as certain of their projections and related fibres and conditional measures. These random sponges can be also viewed as limit sets of some percolation process on the unit cube endowed with an $(m_1, \ldots, m_k)$-adic grid, where $m_1 \geq \cdots \geq m_k \geq 2$ are integers.

Until now the Hausdorff dimension of such a set K is known in the deterministic case and only when $k = 2$ in the random case, while the projections of K and the random Mandelbrot measures to be considered have been studied in the conformal case only. Understanding the missing cases is our goal, with Mandelbrot measures and their projections as a main tool for a variational approach to the Hausdorff dimension of K and its orthogonal projections, together with a "weighted" version of the thermodynamic formalism on symbolic spaces.

Before coming in more details to these objects and our motivations, it seems worth giving an overview of the nature of the main results known in dimension theory of self-affine sets. Recall that given an integer $N \geq 2$ and an iterated function system (IFS) $\{f_i\}_{1 \leq i \leq N}$ of contractive maps of a complete metric space X, there exists a unique non-empty compact set $K \subset X$ such that

$$K = \bigcup_{i=1}^{N} f_i(K),$$

(see [32]). When X is a Euclidean space and f_i are affine maps, due to the above equality K is called self-affine. In particular, K is called self-similar if f_i are all similitudes; we will not focus on the self-similar case and refer the reader to [30] for a recent survey of this topic.

The first class of strictly self-affine sets that have been studied in detail are certainly Bedford-McMullen carpets in $\mathbb{R}^2$, also known as self-affine Sierpinski carpets. They are among the most natural classes of fractal sets having different Hausdorff and box counting dimensions. To be specific, one fixes two integers $m_1 > m_2 \geq 2$ and a subset $A \subset \{0, \ldots, m_1-1\} \times \{0, \ldots, m_2-1\}$ of cardinality at least 2; the Bedford-McMullen carpet K associated with A is the attractor of the system

$$S_A = \left\{ f_a(x_1, x_2) = \left(\frac{x_1 + a_1}{m_1}, \frac{x_2 + a_2}{m_2} \right) : a = (a_1, a_2) \in A \right\}$$

of contractive affine maps of the Euclidean plane; note that by construction $K \subset [0,1]^2$. Set $N_i = \#\{(a_1, a_2) \in A : a_2 = i\}$ for all $0 \leq i \leq m_2 - 1$ and $\psi : \theta \in \mathbb{R}_+ \mapsto \log \sum_{i=0}^{m_2-1} N_i^\theta$. Bedford and McMullen proved independently [8, 44] that

$$\dim_H K = \frac{\psi(\alpha)}{\log(m_2)}, \text{ where } \alpha = \frac{\log(m_2)}{\log(m_1)},$$

and

$$\dim_B K = \frac{\psi(1)}{\log(m_1)} + \left(\frac{1}{\log(m_2)} - \frac{1}{\log(m_1)} \right) \psi(0),$$

where $\dim_H$ and $\dim_B$ respectively stand for the Hausdorff and the box counting dimension, and $\psi(1)$ and $\psi(0)$ are the topological entropy of K and that of its projection to the x_2-axis respectively. Moreover, $\dim_H K = \dim_B K$ if and only if the positive N_i are all equal. Note that the possible dimension gap $\dim_H K < \dim_B K$ cannot hold for self-similar sets (see [18]). For a general self-affine Sierpinski sponge $K \subset [0,1]^k$ invariant under the action of an expanding diagonal endomorphism f of $\mathbb{T}^k$ with eigenvalues the integers $m_1 > \cdots > m_k \geq 2$ (we identify $[0,1]^k$ with $\mathbb{T}^k$), similar formulas as in the 2-dimensional case hold. In particular, the approach to dimensional properties of compact f-invariant sets developed by Kenyon and Peres [38] extends the results of [44] and establishes the following variational principle: the Hausdorff dimension of K is the supremum of the Hausdorff dimensions of ergodic measures supported on K, i.e. the Hausdorff and

dynamical dimensions of K coincide. Moreover, the dimension of any ergodic measure is given by the Ledrappier-Young formula

$$(1.1) \qquad \dim(\mu) = \frac{1}{\log(m_1)} \cdot h_\mu(f) + \sum_{i=2}^{k} \Big(\frac{1}{\log(m_i)} - \frac{1}{\log(m_{i-1})} \Big) h_{\mu \circ \Pi_i^{-1}}(\Pi_i \circ f),$$

where $\Pi_i : \mathbb{R}^k \to \mathbb{R}^{k-i+1}$ is the projection defined by $(x_1, \ldots, x_k) \mapsto (x_i, \ldots, x_k)$. Also, in the variational principle, the supremum is uniquely attained at some Bernoulli product measure.

Dimension theory of self-affine sets has been developed to understand the case of more "generic" self-affine IFSs as well. First, given a family $\{M_i\}_{1 \le i \le N}$ of linear automorphisms of $\mathbb{R}^k$ to itself whose norm subordinated to the Euclidean norm on $\mathbb{R}^k$ is smaller than $1/2$, it was shown [17, 34] that for $\mathcal{L}^{Nk}$-almost every choice of N translation vectors $v_1, \ldots, v_N$ in $\mathbb{R}^k$, the Hausdorff dimension of the attractor K of the affine IFS $\{M_i x + v_i\}_{1 \le i \le N}$ is the maximum of the Hausdorff dimensions of the natural projections of ergodic measures on $(\{1, \ldots, N\}^{\mathbb{N}^*}, \sigma)$ to K (here σ stands for the left shift operation); in addition for such a measure μ, $\dim(\mu) = \min(k, \dim_L(\mu))$, where $\dim_L(\mu)$ is the so-called Lyapunov dimension of μ [17, 48, 34, 33]. A similar result holds for $\mathcal{L}^{Nk^2}$-almost every choice of contractive $\{M_i\}_{1 \le i \le N}$ in some non-empty open set, with a fixed $(v_1, \ldots, v_N)$ [2]. In these contexts one has $\dim_H K = \dim_B K = \min(k, \dim_A K)$, where $\dim_A K$ stands for the so-called affinity dimension of K (note that for a Bedford-McMullen carpet, the affinity dimension equals $\frac{\psi(1)}{\log(m_2)}$ if $\psi(1) \le \log(m_2)$ and $\frac{\psi(1)}{\log(m_1)} + \big(\frac{1}{\log(m_2)} - \frac{1}{\log(m_1)} \big) \log(m_2)$ otherwise, so that in this case the previous equality between dimensions only occurs exceptionally). If both $\{M_i\}_{1 \le i \le N}$ and $(v_1, \ldots, v_N)$ are fixed, the following stronger result appeared recently in the 2-dimensional case: if $\{M_i\}_{1 \le i \le N}$ satisfies the strong irreducibility property and $\{M_i / \sqrt{|\det(M_i)|}\}_{1 \le i \le N}$ generate a non-compact group in $GL_2(\mathbb{R})$, and if the IFS $\{f_i\}_{1 \le i \le N}$ is exponentially separated, then $\dim_H K$ is the supremum of the Lyapounov dimensions of the self-affine measures supported on K (it is not known whether this supremum is reached in general); here again for such a measure, the Lyapunov and Hausdorff dimensions coincide, and $\dim_H K = \dim_B K = \dim_A K$ [3, 31]. The last two contexts make a central use of the notion of Furstenberg measure associated to a self-affine measure, whose crucial role in the subject was first pointed out in [21]. There are also similar results in the case that the strong irreducibility fails but the M_i cannot be simultaneously reduced to diagonal automorphisms [22, 4, 3, 31].

Let us come back to self-affine carpets. Their study was further developed with the introduction of Gatzouras-Lalley carpets [41] (with an application to the study of some non-conformal nonlinear repellers [28]) and Baranski carpets [1]. There, the linear parts are no more subject to be equal, but they are still diagonal, and it is not true in general that there is a unique ergodic measure with maximal Hausdorff dimension [35, 5] (see also [40] for a study of Gatzouras-Lalley type carpets when the linear parts are trigonal). It turns out that extending the dimension theory of these carpets to the higher dimensional case raises serious difficulties in general, as it was shown in [13] that the attractor may have a Hausdorff dimension strictly larger than its dynamical dimension.

On the side of random fractal sets, one naturally meets random statistically self-affine sets. Such a set K obeys almost surely an equation of the form $K(\omega) = \bigcup_{i=0}^{N} f_i^\omega(K_i(\omega))$, where the f_i^ω are random contractive affine maps and the sets K_i are copies of K. Results similar to those obtained for almost all self-affine sets described above exist in the following

situation: the sets K_i are mutually independent and independent of the f_i, the linear maps of the f_i are deterministic, but the translation parts are i.i.d and follow a law compactly supported and absolutely continuous with respect to $\mathcal{L}^d$ [33]. Results are also known for random Sierpinski carpets. There are two natural ways to get such random sets. The first one falls in the setting of random dynamical systems. It consists in considering an ergodic dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, T)$, on which is defined a random non-empty subset $A(\omega)$ of $\{0, \ldots, m_1 - 1\} \times \{0, \ldots, m_2 - 1\}$ such that $\mathbb{E}(\#A) > 1$. Then one starts with the set of maps $S_{A(\omega)}$, and recursively, at each step $n \geq 2$ of the iterative construction of the random attractor $K(\omega)$, replace the set of contractions $S_{A(T^{n-2}(\omega))}$ by $S_{A(T^{n-1}(\omega))}$, so that the contractive maps used after n iterations take the form $f_{a_0} \circ \cdots \circ f_{a_{n-1}}$, with $a_i \in S_{A(T^i(\omega))}$. By construction, $K(\omega) = \bigcup_{a \in A(\omega)} f_a(K(\sigma(\omega)))$. The Hausdorff and box-counting dimensions of such sets and their higher dimensional versions have been determined in [39] (in a slightly more general setting); the situation is close to that in the deterministic one.

The other natural way to produce random statistically self-affine carpets is related to branching processes and consists in using a general percolation scheme detailed below; at the moment let us just say that one starts with a possibly empty random subset $A(\omega)$ of $\{0, \ldots, m_1 - 1\} \times \{0, \ldots, m_2 - 1\}$, and again assumes that $\mathbb{E}(\#A) > 1$. Then, one constructs on the same probability space the set $A(\omega)$ and a random compact set $K(\omega) \subset \mathbb{T}^2$, and $m_1 \times m_2$ random compact sets $K(a, \omega)$, $a \in \{0, \ldots, m_1 - 1\} \times \{0, \ldots, m_2 - 1\}$, so that $K(\omega) = \bigcup_{a \in A(\omega)} f_a(K(a, \omega))$, where the $K(a)$ are independent copies of K, and they are also independent of A. The set K is non-empty with positive probability. These random sets have been studied in [27], and their self-similar versions have been investigated extensively (see e.g. [29, 50, 47]). Setting now $\psi(\theta) = \log \sum_{i=0}^{m_2-1} \mathbb{E}(N_i)^\theta$ and letting t be the unique point at which the convex function ψ attains its minimum over $[0, 1]$ if ψ is not constant, and $t = 1$ otherwise, one has, with probability 1, conditional on $K \neq \emptyset$,

$$(1.2) \qquad \dim_H K = \frac{\psi(\alpha)}{\log(m_2)} \text{ where } \alpha = \max\left(t, \frac{\log(m_2)}{\log(m_1)}\right),$$

and

$$\dim_B K = \frac{\psi(1)}{\log(m_1)} + \left(\frac{1}{\log(m_2)} - \frac{1}{\log(m_1)}\right)\psi(t).$$

Moreover, $\dim_H K = \dim_B K$ iff $t = 1$ or all the positive $\mathbb{E}(N_i)$ are equal (we note that the value of $\dim_H K$ was previously obtained in [46, 9, 10] in the very special case that there is an integer $b \geq 2$ such that the law of A assigns equal probabilities to subsets of cardinality b and probability 0 to the other ones). It is worth pointing out that the origin of this different formula with respect to the deterministic case comes from the possibility that $\mathbb{E}(N_i) < 1$ for some i, which makes the situation quite versatile with respect to the deterministic Bedford-McMullen carpets.

The approach developed in [27] to get (1.2) is not based on a variational principle related to a natural class of measures supported on the attractor. To determine the sharp lower and upper bounds for $\dim_H K$, the authors of [27] adapted the approach used by Bedford in the deterministic case: the lower bound for $\dim_H K$ is obtained by taking the maximum of two values, namely $\dim_H \Pi_2(K)$ (where Π_2 still denotes the orthogonal projection on the x_2-axis) and the maximum of the lower bounds for the Hausdorff dimensions of certain random subsets of K. Each such subset E is obtained by considering the union of almost all the fibres $\Pi_2^{-1}(\{(0, x_2)\}$ with respect to the restriction to $\Pi_2(K)$ of some Bernoulli product measure. The Hausdorff dimensions of $\Pi_2(E)$ and that of the associated fibres

4

are controlled from bellow. This yields a lower bound for $\dim_H E$ thanks to a theorem of Marstrand. The upper bound for $\dim_H K$ is obtained by using some effective coverings of K. It turns out to be delicate to transfer these methods to the higher dimensional cases. Indeed, for the lower bound, the Hausdorff dimension of the 1-dimfibreensional fibres mentioned above is obtained thanks to statistically self-similar branching measures in random environment, the dimension of which is relatively direct to get, and it yields the dimension of the fibre. Using this approach in the higher dimensional case $k \geq 3$, we would have to consider the restriction of Bernoulli measures to $\Pi_i(K)$ for $2 \leq i \leq k$. Then for $i \geq 3$ we would meet the much harder problem to estimate the Hausdorff dimension of fibres which are statistically self-affine Sierpinski sponges in a random environment in $\mathbb{R}^{i-1}$, a problem not less difficult than the one we consider in this paper; one would have to compute $\dim_H \Pi_i(K)$ as well, a question that we will naturally consider. Also, for the upper bound, extending to higher dimensions the combinatorial argument used in [27] to get effective coverings seems impossible. However, we will see that it is rather direct to get the box-counting dimension of K in the higher dimensional cases by adapting the arguments of [27] used for the two dimensional case.

We will develop the dimension theory for statistically self-affine Sierpinski sponges in $\mathbb{R}^k$, for any $k \geq 2$, by studying the statistiscally self-affine measures on K (which are also called Mandelbrot measures on K). We will prove the following Ledrappier-Young type formula: given a Mandelbrot measure μ on K (see Sections 2.2 and 2.6 for the definition),

$$\dim(\mu) = \frac{1}{\log(m_1)} \dim_e(\mu) + \sum_{i=2}^{k} \Big(\frac{1}{\log(m_i)} - \frac{1}{\log(m_{i-1})} \Big) \dim_e(\mu \circ \Pi_i^{-1})$$

$$(1.3) \qquad = \frac{1}{\log(m_1)} \dim_e(\mu) + \Big(\frac{1}{\log(m_i)} - \frac{1}{\log(m_{i-1})} \Big) \cdot \min\big(\dim_e(\mu), h_{\nu_i}(\Pi_i \circ f) \big),$$

where $\dim_e(\mu)$ is the entropy dimension of μ, and ν_i is the Bernoulli product measure $\mathbb{E}(\mu \circ \Pi_i^{-1})$ (see Theorem 2.2). The fact that $\dim_e(\mu \circ \Pi_i^{-1}) = \min(\dim_e(\mu), h_{\nu_i}(\Pi_i \circ f))$ follows from our previous study of projections of Mandelbrot measures in [7]. To get (1.3), we show that τ_μ, the L^q-spectrum of μ, is differentiable at 1 with $\tau'_\mu(1)$ equal to the right hand side of (1.3); this implies the exact dimensionality of μ, with dimension equal to $\tau'_\mu(1)$.

Optimising (1.3) yields the sharp lower bound for the Hausdorff dimension of K (see Theorem 4.5); the supremum is uniquely attained, and the optimisation problem is non-standard; the presence of the $k-1$ minima in the sum gives rise to k possible simplifications of the formula separated by what can be thought of as $k-1$ phase transitions according to the position of $\dim_e(\mu)$ with respect to the entropies $h_{\nu_i}(\Pi_i \circ f)$, hence k distinct optimisation problems must be considered, of which the optima must be compared. This study will use the thermodynamic formalism as well as a version of von Neumann's min-max theorem, and the optimal Hausdorff dimension will be expressed as the "weighted" pressure of some deterministic potential (see Theorem 2.3). Our sharp upper bound for $\dim_H K$ proves that this maximal Hausdorff dimension of a Mandelbrot measure supported on K yields $\dim_H K$. This bound is derived from a variational principle as well, namely we optimise over uncountably many types of coverings of K, each of which provides an upper bound for $\dim_H K$ (see Theorem 4.10); when $k = 2$, this does not reduce back to the argument developed in [27]. As a by-product, we get an alternative approach to the proof by Kenyon and Peres [38] of the sharp upper bound in the deterministic higher dimensional

case. One may wonder if the approach by Kenyon and Peres, which in the deterministic case uses a uniform control of the lower local dimension of the unique Bernoulli measure of maximal Hausdorff dimension on K, can be extended to the random case by using the unique Mandelbrot measure of maximal Hausdorff dimension on K. We met an essential difficulty in trying to follow this direction, except in special cases (see the discussion at the beginning of Section 4.4).

Our result regarding the box-counting dimension of K is stated in Theorem 2.4. Another difference between the deterministic or random dynamical Sierpinski sponges, and the random Sierpinski sponges studied in this paper is that for the first two classes of objects, their images by the natural projections Π_i, $2 \leq i \leq k$, are lower dimensional objects of the same type (e.g. Sierpinski sponges project to 2-dimensional Sierpinski carpets via Π_{k-1}), while this is not the case for the third one. Our approach will also provide $\dim_H \Pi_i(K)$ (via a variational principle) and $\dim_B \Pi_i(K)$ for the third class of random attractors (see Section 2.4; we note that the case of $\Pi_k(K)$ reduces to that of $\Pi_2(K)$ when $k = 2$ and K is a statistically self-similar set, a situation which is covered by [15, 19]). Finally, we determine the dimension of the conditional measures associated with the successive images of any Mandelbrot measure by the projections Π_i (Section 2.5), which for each i equals $\dim(\mu) - \dim(\Pi_i\mu)$, hence the conservation dimension formula holds.

Since using symbolic spaces to encode the Euclidean situation is necessary, we will work on such spaces and endow them with adapted ultrametrics, so that the case of random statistically self-affine Euclidean sponges and their projections will be reducible to a particular situation of a more general framework on symbolic spaces and their factors.

Our framework and main results are presented in the next section.

2. Main results on self-affine symbolic spaces, and application to the Euclidean case

We start with defining the symbolic random statistically self-affine sponges, which will be studied in this paper.

2.1. Symbolic random statistically self-affine sponges. Let us first recall the notion of self-affine symbolic space.

Let $\mathbb{N}$, $\mathbb{N}^*$, $\mathbb{R}_+$ and $\mathbb{R}_+^*$ stand for the sets of non-negative integers, positive integers, non-negatice real numbers, and positive real numbers respectively.

Let $k \geq 2$ be an integer. Assume that (X_i, T_i) $(i = 1, \ldots, k)$ are one-sided full-shift spaces over finite alphabets $\mathcal{A}_i$ of cardinality ≥ 2 and such that X_{i+1} is a factor of X_i with a one-block factor map $\pi_i : X_i \to X_{i+1}$ for $i = 1, \ldots, k-1$. That is, π_i is extended from a map (which is also denoted by π_i for convenience) from $\mathcal{A}_i$ to $\mathcal{A}_{i+1}$ by

$$\pi_i \left((x_n)_{n=1}^\infty\right) = (\pi_i(x_n))_{n=1}^\infty \text{ for all } (x_n)_{n=1}^\infty \in X_i.$$

For convenience, we use π_0 to denote the identity map on X_1. Define $\Pi_i : X_1 \to X_i$ by

$$\Pi_i = \pi_{i-1} \circ \pi_{i-2} \circ \cdots \circ \pi_0$$

for $i = 1, \ldots, k$. Define $\mathcal{A}_i^* = \bigcup_{n \geq 0} \mathcal{A}_i^n$, where $\mathcal{A}_i^0$ consists of the empty word ϵ. The maps π_i and Π_i naturally extend to $\mathcal{A}_i^*$ and $\mathcal{A}_1^*$ respectively, that is,

$$\pi_i(x_1 \cdots x_n) = \pi_i(x_1) \cdots \pi_i(x_n), \quad \Pi_i(x_1 \cdots x_n) = \Pi_i(x_1) \cdots \Pi_i(x_n).$$

If $u \in \mathcal{A}_i^*$, we denote by $[u]$ the cylinder made of those elements in X_i which have u as prefix. If $x \in X_i$ and $n \geq 0$, we denote by $x_{|n}$ the prefix of x of length n. For $x, y \in X_i$, let $x \wedge y$ denote the longest common prefix of x and y.

Let $\vec{\gamma} = (\gamma_1, \ldots, \gamma_k) \in \mathbb{R}_+^* \times (\mathbb{R}_+)^{k-1}$. Define an ultrametric distance $d_{\vec{\gamma}}$ on X_1 by

$$(2.1) \qquad d_{\vec{\gamma}}(x, y) = \max \left(e^{-\frac{|\Pi_i(x) \wedge \Pi_i(y)|}{\gamma_1 + \cdots + \gamma_i}} \ : \ 1 \leq i \leq k \right),$$

where

$$|u \wedge v| = \begin{cases} 0, & \text{if } u_1 \neq v_1, \\ \max\{n : \ u_j = v_j \text{ for } 1 \leq j \leq n\} & \text{if } u_1 = v_1 \end{cases}$$

for $u = (u_j)_{j=1}^{\infty}, v = (v_j)_{j=1}^{\infty} \in X_i$.

The metric space $(X_1, d_{\vec{\gamma}})$ is called a self-affine symbolic space. It is a natural model used to characterize the geometry of compact invariant sets on the k-torus under a diagonal endomorphism [8, 44, 38, 6].

If f is a measurable mapping between a measured space $(X, \mathcal{T}, \nu)$ and a measurable space $(Y, \mathcal{S})$, the image $\nu \circ f^{-1}$ of ν will be simply denote by $f\nu$.

Now, we can define symbolic random statistically self-affine sponges. Let $A = (c_a)_{a \in \mathcal{A}_1}$ be a random variable taking values in $\{0, 1\}^{\mathcal{A}_1}$. It encodes a random subset of $\mathcal{A}_1$, namely $\{a \in \mathcal{A}_1 : \ c_a = 1\}$, which we identify with A. Suppose that $\mathbb{E}(\sum_{a \in \mathcal{A}_1} c_a) > 1$. Let $(A(u))_{u \in \mathcal{A}_1^*}$ be a sequence of independent copies of A. For all $n \in \mathbb{N}$, let

$$K_n = \{x \in X_1 : \ c_{x_i}(x_{|i-1}) = 1 \text{ for all } 1 \leq i \leq n\}$$

$$(2.2) \qquad = \bigcup_{u \in \mathcal{A}_1^n: \ \prod_{i=1}^n c_{u_i}(u_{|i-1})=1} [u].$$

Due to our assumption on A, with positive probability, the set

$$K = \bigcap_{n \geq 1} K_n,$$

is the boundary of a non-degenerate Galton Watson tree with offspring distribution given by that of the random integer $\sum_{a \in \mathcal{A}_1} c_a$. The set K satisfies the following symbolic statistically self-affine invariance property:

$$K = \bigcup_{a \in A} a \cdot K^a,$$

where

$$(2.3) \qquad K^a = \bigcap_{n \geq 1} \left(K_n^a = \bigcup_{u \in \mathcal{A}_1^n: \ \prod_{i=1}^n c_{u_i}(au_{|i-1})=1} [u] \right).$$

We call K a symbolic statistically self-affine Sierpinski sponge. The link with the Euclidean case will be made in Section 2.6.

Next we recall the definition and basic properties of Mandelbrot measures on K, and state our result on their exact dimensionality.

2.2. Mandelbrot measures on K.

These measures will play an essential role in finding a sharp lower bound for $\dim_H K$. Let $W = (W_a)_{a \in \mathcal{A}_1}$ be a non-negative random vector defined simultaneously with A and such that $\{W_a > 0\} \subset \{c_a = 1\}$ for all $a \in \mathcal{A}_1$. Let

$$T(q) = T_W(q) = -\log \mathbb{E} \sum_{a \in \mathcal{A}_1} W_a^q, \quad q \geq 0,$$

and suppose that $T(1) = 0$, i.e. $\mathbb{E}(\sum_{a \in \mathcal{A}_1} W_a) = 1$. Let $((A(u), W(u))_{u \in \mathcal{A}_1^*}$ be a sequence of independent copies of (A, W).

For each $u, v \in \mathcal{A}_1^*$ let

$$Q^u(v) = \prod_{k=1}^{|v|} W_{v_k}(u \cdot v_{|k-1}),$$

and simply denote $Q^\epsilon(v)$ by $Q(v)$. Due to our assumptions on W, for each $u \in \mathcal{A}_1^*$, setting $Y_n(u) = \sum_{|v|=n} Q^u(v)$ the sequence

$$(Y_n(u), \sigma(W_a(uv) : a \in \mathcal{A}_1, v \in \mathcal{A}_1^n))_{n \geq 1}$$

is a non-negative martingale. Denote its limit almost sure limit by $Y(u)$. The mapping

$$\mu : [u] \mapsto Q(u)Y(u)$$

defined on the set of cylinders $\{[u] : u \in \mathcal{A}_1^*\}$ extends to a unique measure μ on $(X_1, \mathcal{B}(X_1))$. This measure was first introduced in [43], and is called a Mandelbrot measure. The support of μ is the set

$$K_\mu = \bigcap_{n \geq 0} \bigcup_{u \in \mathcal{A}_1^n : \ Q(u) > 0} [u] \subset K,$$

where the inclusion $K_\mu \subset K$ follows from the assumption

$$\{W_a(u) > 0\} \subset \{c_a(u) = 1\}$$

for all $u \in \mathcal{A}_1^*$ and all $a \in \mathcal{A}_1$. Moreover, if $T'(1-) > 0$, then $(Y_n(u))_{n \geq 1}$ is uniformly integrable (see [37, 11, 16]), and in this case K_μ is a symbolic statistically self-affine set as well. Also, $K_\mu = K$ almost surely if and only if $\{c_a(u) = 1\} \setminus \{W_a > 0\}$ has probability 0 (see [7, Proposition A.1]). If $T'(1-) \leq 0$, either $\mu = 0$ almost surely (in such case one says that μ is degenerate), or

$$\mathbb{P}(\exists \, a \in \mathcal{A}_1 \text{ such that } W_a = 1 \text{ and } W_{a'} = 0 \text{ for } a' \neq a) = 1$$

(see [37, 11, 16] as well), and in this later case $T'(1-) = 0$, K_μ is a singleton and μ is a Dirac measure [16].

The measure μ is also the weak-star limit of the sequence $(\mu_n)_{n \geq 1}$ defined by distributing uniformly (with respect to the uniform measure on X_1) the mass $Q(u)$ on each $u \in \mathcal{A}_1^n$. It is statistically self-affine in the sense that

$$\mu(B) = \sum_{a \in \mathcal{A}_1} W_a \, \mu^a(\sigma(B \cap [a]))$$

for every Borel set $B \subset X_1$, where μ^a is the copy of μ constructed on K^a with the weights $(W(au))_{u \in \mathcal{A}_1^*}$.

We will make a systematic use of the notion of entropy dimension of measures on X_i. Given a finite Borel measure ν on X_i, the entropy dimension of ν is defined as

$$(2.4) \qquad \dim_e(\nu) = \lim_{n \to \infty} -\frac{1}{n} \sum_{u \in \mathcal{A}_i^n} \nu([u]) \log \nu([u]),$$

whenever the limit exists. If ν is T_i-invariant, one has $\dim_e(\nu) = h_\nu(T_i)$, the entropy of ν with respect to T_i (see e.g. [12]).

Due to the results by Kahane and Peyrière [37, 36], when a Mandelbrot measure μ is non-degenerate (that is, when $\mathbb{P}(\mu \neq 0) > 0$), with probability 1, conditional on $\mu \neq 0$, one has

$$\lim_{n \to \infty} \frac{\log(\mu([x_{|n}]))}{-n} = T'(1-) = -\sum_{a \in \mathcal{A}_1} \mathbb{E}(W_a \log W_a), \quad \text{for } \mu\text{-a.e. } x.$$

It then follows that $\dim_e(\mu)$ exists [23] and $\dim_e(\mu) = T'(1-)$. In particular,

$$(2.5) \qquad \dim_e(\mu) \leq \log \mathbb{E}(\#A),$$

with equality if and only if μ is the so called branching measure, i.e. it is obtained from the random vector $(\mathbf{1}_A(a)/\mathbb{E}(\#A))_{a \in \mathcal{A}_1}$ (see Remark 4.1 for a justification of this uniqueness).

Before stating our first result, we present a result deduced from [7] about the entropy dimension of $\Pi_i \mu$ (in [7] we proved the exact dimensionality of the projection of a Mandelbrot measure on $X_1 \times X_1$ to the first factor X_1, in the case that X_1 is endowed with $d_{\vec{\gamma}}$, where $\vec{\gamma} = ((\log \#\mathcal{A}_1)^{-1})$; but projecting from X_1 to X_i and considering the entropy dimension do not affect the arguments):

Theorem 2.1. [7, Theorem 3.2] *Let μ be a non-degenerate Mandelbrot measure on K. Suppose that $T(q) > -\infty$ for some $q > 1$. With probability 1, conditional on $\mu \neq 0$, for all $2 \leq i \leq k$, one has $\dim_e(\Pi_i \mu) = \min(\dim_e(\mu), h_{\nu_i}(T_i))$, where ν_i is the Bernoulli product measure on X_i obtained as $\nu_i = \mathbb{E}(\Pi_i \mu)$.*

Hence, $\dim_e(\mu)$ and $h_{\nu_i}(T_i)$ compete in the determination of the entropy dimension of the ith projection of μ.

Recall that a locally finite Borel measure ν on a metric space (X, d) is said to be exact dimensional with dimension D if

$$\lim_{r \to 0+} \frac{\log(\nu(B(x, r)))}{\log(r)} = D$$

for ν-almost every x. We denote the number D by $\dim(\nu)$ and call it the dimension of ν. Our result on the exact dimensionality of the Mandelbrot measure μ on $(X_1, d_{\vec{\gamma}})$ is the following.

Theorem 2.2 (Exact dimensionality of μ). *Let μ be a non-degenerate Mandelbrot measure on K. Suppose that $T(q) > -\infty$ for some $q > 1$. With probability 1, conditional on $\mu \neq 0$, the measure μ is exact dimensional and $\dim(\mu) = \dim_e^{\vec{\gamma}}(\mu)$, where*

$$(2.6) \qquad \dim_e^{\vec{\gamma}}(\mu) := \sum_{i=1}^k \gamma_i \dim_e(\Pi_i \mu) = \gamma_1 \dim_e(\mu) + \sum_{i=2}^k \gamma_i \min(\dim_e(\mu), h_{\nu_i}(T_i)).$$

This result will follow from the stronger fact that the L^q-spectrum of μ is differentiable at 1 with derivative $\dim_e^{\vec{\gamma}}(\mu)$ (see Theorem 3.1).

2.3. The Hausdorff and box counting dimensions of K.

To state our result on $\dim_H K$, we need to recall some elements of the weighted thermodynamic formalism.

For $1 \leq i \leq k$ and $b \in \mathcal{A}_i = \Pi_i(\mathcal{A}_1)$, let

(2.7) $$N_b^{(i)} = \#\{a \in \mathcal{A}_1 : [a] \subset \Pi_i^{-1}([b]) \text{ and } [a] \cap K \neq \emptyset\}.$$

Then set

$$\widetilde{\mathcal{A}}_i = \{b \in \mathcal{A}_i : \mathbb{E}(N_b^{(i)}) > 0\}.$$

Without loss of generality we assume that $\#\widetilde{\mathcal{A}}_k \geq 2$. Indeed, if $\widetilde{\mathcal{A}}_k$ is a singleton, then X_k plays no role in the geometry of K since $\Pi_k(K)$ is a singleton when $K \neq \emptyset$. As a consequence, $\#\widetilde{\mathcal{A}}_i \geq 2$ for all $2 \leq i \leq k$.

For $1 \leq i \leq k$, let $\widetilde{X}_i$ denote the one-sided symbolic space over the alphabet $\widetilde{\mathcal{A}}_i$. If $\phi : \widetilde{X}_i \to \mathbb{R}$ is a continuous function on $\widetilde{X}_i$, $\vec{\beta} = (\beta_i, \beta_{i+1}, \ldots, \beta_k) \in \mathbb{R}_+^* \times \mathbb{R}_+^{k-i}$, and $\nu \in \mathcal{M}(X_i, T_i)$, let

(2.8) $$h_\nu^{\vec{\beta}}(T_i) := \sum_{j=i}^k \beta_j h_{\Pi_{i,j}\nu}(T_j),$$

where

(2.9) $$\Pi_{i,j} = \pi_{j-1} \circ \cdots \circ \pi_i \quad \text{if } j > i$$

and $\Pi_{i,i}$ is the identity map of X_i, and define the weighted pressure function

(2.10) $$P^{\vec{\beta}}(\phi, T_i) = \sup\left\{\nu(\phi) + h_\nu^{\vec{\beta}}(T_i) : \nu \in \mathcal{M}(\widetilde{X}_i, T_i)\right\},$$

where $\nu(\phi) = \int_{\widetilde{X}_i} \phi \, d\nu$. It is known ([6]) that if ϕ is Hölder-continuous, then the supremum in defining $P^{\vec{\beta}}(\phi, T_i)$ is reached at a unique fully supported measure ν_ϕ, called the equilibrium state of ϕ for $P^{\vec{\beta}}(\phi, T_i)$. Moreover, the mapping $\theta \mapsto P^{\vec{\beta}}(\theta\phi, T_i)$ is differentiable, and

(2.11) $$\frac{d P^{\vec{\beta}}(\theta\phi, T_i))}{d\theta} = \int_{\widetilde{X}_i} \phi(x) \, d\nu_{\theta\phi}(x) = \nu_{\theta\phi}(\phi).$$

For $2 \leq i \leq k$ let $\vec{\gamma}^i = (\vec{\gamma}_j^i)_{i \leq j \leq k} = (\gamma_1 + \cdots + \gamma_i, \gamma_{i+1}, \ldots, \gamma_k)$ and let ϕ_i be the Hölder-continuous potential defined on $\widetilde{X}_i$ by

(2.12) $$\phi_i(x) = (\gamma_1 + \cdots + \gamma_i) \log \mathbb{E}(N_{x_1}^{(i)}).$$

For this potential and $\vec{\beta} = \vec{\gamma}^i$, we set

(2.13) $$P_i(\theta) = P^{\vec{\gamma}^i}(\theta\phi_i, T_i), \quad \theta \in \mathbb{R}.$$

We also define $P_{k+1} = P_k$.

The equilibrium state of $\theta\phi_i$ for $P^{\vec{\gamma}^i}$ will be also called the equilibrium state of P_i at θ.

By (2.11), we have

(2.14) $$P_i'(\theta) = (\gamma_1 + \cdots + \gamma_i) \sum_{b \in \widetilde{\mathcal{A}}_i} \nu_{\theta\phi_i}([b]) \log \mathbb{E}(N_b^{(i)}) \quad (\theta \in \mathbb{R}).$$

10

We now define some parameters involved in the next statement. In order to slightly simplify the exposition, we assume that all the γ_i are positive. The general situation will be considered in Section 8.

Set $I = \{2, \ldots, k\}$ (introducing this convention will simplify the discussion in Section 8),

$$(2.15) \qquad \widetilde{\theta}_i = \frac{\gamma_1 + \cdots + \gamma_{i-1}}{\gamma_1 + \cdots + \gamma_i} \quad \text{if } 2 \leq i \leq k,$$

and define

$$\widetilde{I} = \left\{ i \in I : P_i'(1) \geq 0 \right\}.$$

Moreover, define

$$(2.16) \qquad i_0 = \begin{cases} \min(\widetilde{I}) & \text{if } \widetilde{I} \neq \emptyset \\ k+1 & \text{if } \widetilde{I} = \emptyset \end{cases}$$

and

$$(2.17) \qquad \theta_{i_0} = \begin{cases} \min \left\{ \theta \in [\widetilde{\theta}_{i_0}, 1] : P_{i_0}'(\theta) \geq 0 \right\} & \text{if } i_0 \leq k \\ 1 & \text{if } i_0 = k+1 \end{cases}.$$

Theorem 2.3 (Hausdorff dimension of K). *With probability 1, conditional on $K \neq \emptyset$,*

$$\dim_H K = \sup\{\dim_e^{\vec{\gamma}}(\mu) : \ \mu \text{ is a positive Mandelbrot measure supported on } K\}$$

$$(2.18) \qquad = \begin{cases} P_{i_0}(\theta_{i_0}) & \text{if } i_0 \leq k \\ P_k(1) & \text{if } i_0 = k+1 \end{cases}$$

$$= \inf \left\{ P_i(\theta) : \ i \in I, \ \widetilde{\theta}_i \leq \theta \leq 1 \right\}.$$

Moreover, there is a unique Mandelbrot measure jointly defined with K at which the supremum is attained, conditional on $K \neq \emptyset$.

The Mandelbrot measure at which the supremum is attained will be specified in Section 4.2. It is constructed from the equilibrium state of P_{i_0} at θ_{i_0}.

Theorem 2.4 (Box counting dimension of K). *With probability 1, conditional on $K \neq \emptyset$,*

$$\dim_B K = \gamma_1 \log \mathbb{E}(\#A) + \sum_{i=2}^{k} \gamma_i \min_{\theta \in [0,1]} \log \sum_{b \in \widetilde{\mathcal{A}}_i} \mathbb{E}(N_b^{(i)})^\theta.$$

Next we give a necessary and sufficient condition for $\dim_H K = \dim_B K$. Define

$$(2.19) \qquad \psi_i : \theta \in [0,1] \mapsto \log \sum_{b \in \widetilde{\mathcal{A}}_i} \mathbb{E}(N_b^{(i)})^\theta \quad (2 \leq i \leq k).$$

For each $2 \leq i \leq k$, denote by $\widehat{\theta}_i$ the point in $[0,1]$ at which ψ_i reaches its minimum if ψ_i is not constant (i.e. there is $b \in \widetilde{\mathcal{A}}_i$ such that $\mathbb{E}(N_b^{(i)}) \neq 1$), and $\widehat{\theta}_i = 0$ otherwise.

We will need the following lemma to state and prove Corollary 2.6 about the necessary and sufficient condition for the equality $\dim_H K = \dim_B K$ to hold.

Lemma 2.5. *For each $2 \leq i \leq k$, ψ_i takes the value $\log \mathbb{E}(\#A)$ at $\theta = 1$. Moreover if $2 \leq i \leq k-1$, then $\psi_i \geq \psi_{i+1}$, and $\widehat{\theta}_i < 1$ implies $\widehat{\theta}_{i+1} < 1$.*

Proof. The first property is due to the relation $\mathbb{E}(N_{\widetilde{b}}^{(i+1)}) = \sum_{b\in\widetilde{\mathscr{A}}_i:\,\pi_i(b)=\widetilde{b}}\mathbb{E}(N_b^{(i)})$ for any $\widetilde{b}\in\widetilde{\mathscr{A}}_{i+1}$, and the second one is due both to this property and the subadditivity of $y\geq 0\mapsto y^\theta$. The third property is a direct consequence of the first two properties. Indeed if $\widehat{\theta}_{i+1}=1$, then by the definition of $\widehat{\theta}_{i+1}$ and the convexity of ψ_{i+1}, we have for every $\theta\in[0,1)$, $\psi_{i+1}(\theta)>\psi_{i+1}(1)=\log\mathbb{E}(\#A)$ and so

$$\psi_i(\theta)\geq\psi_{i+1}(\theta)>\log\mathbb{E}(\#A)=\psi_i(1),$$

which implies that $\widehat{\theta}_i=1$. $\qquad\square$

Corollary 2.6. *It holds that* $\dim_H K = \dim_B K$ *with probability 1, conditional on* $K\neq\emptyset$, *if and only if* $\mathbb{E}(N_b^{(i)})$ *does not depend on* $b\in\widetilde{\mathscr{A}}_i$ *for all* $i\in I$ *such that* $\widehat{\theta}_i<1$.

2.4. **Dimensions of projections of μ and K.** We still assume that all the γ_i are positive and will discuss the general case in Section 8. In the next statement μ is assumed to be a non-degenerate Mandelbrot measure such that $K_\mu\subset K$ almost surely.

Theorem 2.7 (Dimension of $\Pi_i\mu$). *Let* $2\leq i\leq k$. *Suppose that* $T(q)>-\infty$ *for some* $q>1$. *With probability 1, conditional on* $\mu\neq 0$, *the measure* $\Pi_i\mu$ *is exact dimensional and* $\dim(\Pi_i\mu)=\dim_e^{\vec{\gamma}^i}(\Pi_i\mu)$, *where*

$$\dim_e^{\vec{\gamma}^i}(\Pi_i\mu):=\sum_{j=i}^{k}\vec{\gamma}_j^i\dim_e(\Pi_j\mu)=\sum_{j=i}^{k}\vec{\gamma}_j^i\min(\dim_e(\mu),h_{\nu_j}(T_j)).$$

Next we consider $\dim_H\Pi_i(K)$. For this purpose, let $i\in\{2,\ldots,k\}$. Define $\overline{\theta}_i^i=0$ and $\overline{\theta}_j^i=\widetilde{\theta}_j$ for $j=i+1,\ldots,k$ (cf. (2.15) for the definition of $\widetilde{\theta}_j$). Set $I_i=\{i,\ldots,k\}$ and define

$$\widetilde{I}_i=\{j\in I_i:\ P_j'(1)\geq 0\}.$$

Then define

$$j_0:=j_0(i)=\begin{cases}\min(\widetilde{I}_i)&\text{if }I_i\neq\emptyset\\ k+1&\text{otherwise}\end{cases}.$$

Also, set

$$(2.20)\qquad\theta_{j_0}^i=\begin{cases}\min\left\{\theta\in[\overline{\theta}_{j_0}^i,1]:P_{j_0}'(\theta)\geq 0\right\}&\text{if }j_0\leq k\\ 1&\text{if }j_0=k+1\end{cases}.$$

Theorem 2.8 (Hausdorff dimension of $\Pi_i(K)$). *Let* $2\leq i\leq k$. *Let* $j_0=j_0(i)$ *and* $\theta_{j_0}^i$ *be defined as above. Then, with probability 1, conditional on* $K\neq\emptyset$,

$$\dim_H\Pi_i(K)=\sup\{\dim_e^{\vec{\gamma}^i}(\Pi_i\mu):\mu\text{ is a positive Mandelbrot measure supported on }K\}$$

$$(2.21)\qquad=\begin{cases}P_{j_0}(\theta_{j_0}^i)&\text{if }j_0\leq k\\ P_{j_0}(1)&\text{if }j_0=k+1\end{cases},$$

and the supremum is uniquely attained at a Mandelbrot measure jointly defined with K if and only if one of the following conditions is fulfilled:

(a) $j_0>i$;
(b) $j_0=i$ and $\theta_{j_0}^i>0$;
(c) $j_0=i$, $\theta_{j_0}^i=0$ and $P_{j_0}'(0)=0$.

12

Remark 2.9. *We deduce from Theorems 2.3 and 2.8 that:*

(i) *if $2 \le i < i_0$ then $\dim_H \Pi_i(K) = \dim_H K$.*

(ii) *$\dim_H K = \dim_H \Pi_{i_0}(K)$ if and only if $i_0 \le k$ and $P'_{i_0}(\theta_{i_0}) = 0$, or if $i_0 = k+1$. Indeed, it is clear that the condition is sufficient. Now suppose that $\dim_H K = \dim_H \Pi_{i_0}(K)$. If $i_0 \le k$ and $P'_{i_0}(\theta_{i_0}) > 0$ then $\theta_{i_0} = \widetilde{\theta}_{i_0} > 0$. Moreover, $j_0(i_0) = i_0$ and $P'_{i_0}(\widetilde{\theta}_{i_0}) > 0$ implies that $\theta_{i_0}^{i_0} < \widetilde{\theta}_{i_0}$, so $\dim_H \Pi_{i_0}(K) = P_{i_0}(\theta_{i_0}^{i_0}) < \dim_H K = P_{i_0}(\widetilde{\theta}_{i_0})$. Thus, if $i_0 \le k$ then $P'_{i_0}(\theta_{i_0}) = 0$.*

(iii) *If $2 \le i < j \le k$, then*

(a) *if $2 \le i < j < j_0(i)$, then $\dim_H \Pi_j(K) = \dim_H \Pi_i(K)$;*

(b) *$\dim_H \Pi_{j_0(i)}(K) = \dim_H \Pi_i(K)$ if and only if $j_0(i) \le k$ and $P'_{j_0(i)}(\theta_{j_0}) = 0$, or $j_0(i) = k+1$. This can be proven like (ii).*

Theorem 2.10 (Box counting dimension of $\Pi_i(K)$). *Let $2 \le i \le k$. With probability 1, conditional on $K \ne \emptyset$,*

$$\dim_B \Pi_i(K) = \sum_{j=i}^{k} \vec{\gamma}_j^i \min_{\theta \in [0,1]} \log \sum_{b \in \widetilde{\mathcal{A}}_j} \mathbb{E}(N_b^{(j)})^\theta.$$

Corollary 2.11. *Let $2 \le i \le k$. It holds that $\dim_H \Pi_i(K) = \dim_B \Pi_i(K)$ with probability 1, conditional on $K \ne \emptyset$, if and only if either of the following three conditions hold:*

(1) *$\widehat{\theta}_i = 1$ and $\mathbb{E}(N_b^{(j)})$ does not depend on $b \in \widetilde{\mathcal{A}}_j$ for all $j \in I_i \setminus \{i\}$ such that $\widehat{\theta}_j < 1$.*

(2) *$0 < \widehat{\theta}_i < 1$, or $\widehat{\theta}_i = 0$ and $\psi'_i(0) = 0$. Moreover setting*

$$j'_0 = \max\{j \in I_i : \text{ for all } j' \le j \text{ in } I_i, \text{ either } 0 < \widehat{\theta}_{j'} < 1, \text{ or } \widehat{\theta}_{j'} = 0 \text{ and } \psi_{j'}(0) = 0\},$$

then the following properties hold:

(i) *For all $j \in I_i$ such that $j \le j'_0$, for all $b \in \widetilde{\mathcal{A}}_j$, $\Pi_{i,j}^{-1}(b) \cap \mathcal{A}_i$ is a singleton (in particular $\widehat{\theta}_j = \widehat{\theta}_i$);*

(ii) *For all $j \in I_i$ such that $j > j'_0$ one has $\widehat{\theta}_j = 0$ and $\psi'_j(0) > 0$, and*

$$\sum_{b' \in \widetilde{\mathcal{A}}_i : [b'] \subset \Pi_{i,j}^{-1}([b])} \mathbb{E}(N_{b'}^{(i)})^{\widehat{\theta}_i}$$

does not depend on $b \in \widetilde{\mathcal{A}}_j$.

(3) *For all $j \in I_i$, $\widehat{\theta}_j = 0$, $\psi'_j(0) > 0$, and $\#\Pi_{i,j}^{-1}(b) \cap \mathcal{A}_i$ does not depend on $b \in \widetilde{\mathcal{A}}_j$.*

In the random case, Theorem 2.8, Theorem 2.10 and Corollary 2.11 are new except in the case $i = k$, which is reducible to the two dimensional case considered in [15, 19, 7].

2.5. **Dimensions of conditional measures.** Given a non-degenerate Mandelbrot measure μ, conditional on $\mu \ne 0$, for each $2 \le i \le k$, the measure μ disintegrates as the skewed product of $\Pi_i \mu(\mathrm{d}z)\, \mu^z(\mathrm{d}x)$, where μ^z is the conditional measure supported on $\Pi_i^{-1}(\{z\}) \cap K$ for $\Pi_i \mu$-almost every z. We will prove the exact dimensionality of the measures μ^z and the value for their dimensions.

Let us start with a consequence of [7, Theorems 3.1 and 3.2]:

13

Theorem 2.12. *Let μ be a non-degenerate Mandelbrot measure supported on K and $2 \leq i \leq k$. Let ν_i be the Bernoulli product measure equal to $\mathbb{E}(\Pi_i\mu)$. With probability 1, conditional on $\mu \neq 0$, $\Pi_i\mu$ is absolutely continuous with respect to ν_i if $\dim_e(\mu) > h_{\nu_i}(T_i)$, otherwise $\Pi_i\mu$ and ν_i are mutually singular. Moreover, if $T(q) > -\infty$ for some $q > 1$, then for $\Pi_i\mu$-a.e. $z \in \Pi_i(K)$ and μ^z-a.e. $x \in K$,*

$$\lim_{n \to \infty} \frac{\log(\mu^z([x_{|n}]))}{-n} = \dim_e(\mu) - \dim_e(\Pi_i\mu),$$

which is equal to $\dim_e(\mu) - h_{\nu_i}(T_i)$ if $\dim_e(\mu) > h_{\nu_i}(T_i)$ and 0 otherwise; in particular the entropy dimension of μ^z exists and equals $\dim_e(\mu) - \dim_e(\Pi_i\mu)$.

It is worth mentioning that the existence of the local entropy dimension for μ^z and the entropy dimension conservation formula comes from the study achieved in [20] for the self-similar case, while the alternative between singularity and absolute continuity regarding $\Pi_i\mu$, as well as the value of $\dim_e(\Pi_i\mu)$ and so that of $\dim_e(\mu^z)$ are obtained in [7].

For the Hausdorff dimension of the conditional measures, we prove the following result:

Theorem 2.13. *Let μ be a non-degenerate Mandelbrot measure supported on K and $2 \leq i \leq k$. Let ν_i be the Bernoulli product measure equal to $\mathbb{E}(\Pi_i\mu)$. Suppose that $T(q) > -\infty$ for some $q > 1$. With probability 1, conditional on $\mu \neq 0$:*

(1) *If $\dim_e(\mu) \leq h_{\nu_i}(T_i)$, then for $\Pi_i\mu$-a.e. $z \in \Pi_i(K)$, the measure μ^z is exact dimensional with Hausdorff dimension equal to 0.*

(2) *If $\dim_e(\mu) > h_{\nu_i}(T_i)$, then for $\Pi_i\mu$-a.e. $z \in \Pi_i(K)$, the measure μ^z is exact dimensional with*

$$(2.22) \qquad \dim(\mu^z) = \gamma_1(\dim_e(\mu) - h_{\nu_i}(T_i)) + \sum_{j=2}^{i-1} \gamma_j\big(\min(\dim_e(\mu), h_{\nu_j}(T_j)) - h_{\nu_i}(T_i)\big).$$

(3) *In both the previous situations, the Hausdorff dimension conservation*

$$\dim(\mu) = \dim(\mu^z) + \dim(\Pi_i\mu)$$

holds.

Naturally, there is a similar result for the conditional measures of $\Pi_i\mu$ projected on X_j, $2 \leq i < j \leq k$.

Theorem 2.14. *Suppose $k \geq 3$. Let μ be a non-degenerate Mandelbrot measure supported on K and $2 \leq i < j \leq k$. Suppose that $T(q) > -\infty$ for some $q > 1$. With probability 1, conditional on $\mu \neq 0$, denote by $(\Pi_i\mu)^{j,z}$ the conditional measure of $\Pi_i\mu$ associated with the projection $\Pi_{i,j}$, and defined for $\Pi_j\mu$-almost every z.*

(1) *If $\dim_e(\mu) \leq h_{\nu_j}(T_j)$, then for $\Pi_j\mu$-a.e. $z \in \Pi_j(K)$, the measure $(\Pi_i\mu)^{j,z}$ is exact dimensional with Hausdorff dimension equal to 0.*

(2) *If $\dim_e(\mu) > h_{\nu_j}(T_j)$, then for $\Pi_j\mu$-a.e. $z \in \Pi_j(K)$, the measure $(\Pi_i\mu)^{j,z}$ is exact dimensional with*

$$(2.23) \qquad \dim((\Pi_i\mu)^{j,z}) = \sum_{j'=i}^{j-1} \vec{\gamma}_{j'}^i\big(\min(\dim_e(\mu), h_{\nu_{j'}}(T_{j'})) - h_{\nu_j}(T_j)\big).$$

14

(3) *In both the previous situations, the Hausdorff dimension conservation*

$$\dim(\Pi_i\mu) = \dim((\Pi_i\mu)^{j,z})) + \dim(\Pi_j\mu)$$

holds for $\Pi_{i,j}$-a.e. $z \in \Pi_{i,j}(K)$.

2.6. Applications to the Euclidean realisations of symbolic random statistically self-affine Sierpinski sponges.

The link with Euclidean random sponges is the following: Given a sequence of integers $2 \leq m_k < \cdots < m_1$, if $\mathcal{A}_i = \prod_{j=i}^{k}\{0,\ldots,m_i-1\}$ for $1 \leq i \leq k$, π_i is the canonical projection from $\mathcal{A}_i$ to $\mathcal{A}_{i+1}$ for $1 \leq i \leq k-1$, $\gamma_1 = 1/\log(m_1)$, and $\gamma_i = 1/\log(m_i) - 1/\log(m_{i-1})$, $2 \leq i \leq k$, then the cylinders of generation n of X_1 project naturally onto parallelepipeds of the family

$$\mathcal{G}_n = \left\{\prod_{i=1}^{k}[\ell_i m_i^{-n}, (\ell_i+1)m_i^{-n}] : 0 \leq \ell_i \leq m_i^n - 1\right\},$$

and K projects on a statistically self-affine Sierpinski sponge $\widetilde{K}$, also called Mandelbrot percolation set associated with $(A(u))_{u \in \mathcal{A}_1^*}$ in the cube $[0,1]^k$ endowed with the nested grids $(\mathcal{G}_n)_{n \geq 0}$.

It is direct to prove that all the results of the previous sections are valid if one replaces K by $\widetilde{K}$, the Mandelbrot measures by their natural projections on $\widetilde{K}$ (which are also called Mandelbrot measures), and Π_i by the orthogonal projection from $\mathbb{R}^k$ to $\{0\}^{i-1} \times \mathbb{R}^{k-i+1}$.

If $\widetilde{K}$ is deterministic, then $i_0 = 2$, $\theta_2 = \gamma_1/(\gamma_1+\gamma_2)$, the Mandelbrot measure of maximal Hausdorff dimension is a Bernoulli product measures, and we recover the result established by Kenyon and Peres in [38] (they work on $(\mathbb{R}/\mathbb{Z})^k$ but it is equivalent); also, in this case the results on the dimension of conditional measures is a special case of the general result obtained by the second author on the dimension theory of self-affine measures [25]. If $k = 2$, the Euclidean version of Theorem 2.2 yields the value of $\dim_H \widetilde{K}$ computed by Gatzouras and Lalley in [27].

Regarding the box counting dimension of $\widetilde{K}$, if $\widetilde{K}$ is deterministic, we just recover the result of [38]; in this case, $\widehat{\theta}_i = 0$ for all $2 \leq i \leq k$. If $k = 2$, we recover the result of Gatzouras and Lalley in [27].

The paper is organized as follows. Section 3 is dedicated to the proof of Theorem 2.2, Section 4 to the proof of Theorem 2.3, Section 5 to that of Theorem 2.4 and its corollary, Section 6 to those of the corresponding results for projections of Mandelbrot measures and K, Section 7 to those on conditional measures, and the brief Section 8 to the case when some γ_i vanish.

3. THE HAUSDORFF DIMENSION OF μ. PROOF OF THEOREM 2.2

Let us start with a few definitions.

With the notation given in the introduction part, for any word $I \in \mathcal{A}_1^*$ and any integer $n \geq 0$, we denote by μ^I the measure defined on X_1 by

$$\mu^I([J]) = Q^I(J)Y(IJ) \quad \text{for } J \in \mathcal{A}_1^*$$

and by μ_n^I the measure on X_1 obtained by distributing uniformly $Q^I(J)$ on any cylinder J of the nth generation. Also, we write $\mu_n = \mu_n^\epsilon$.

For $1 \leq i \leq k$ and $n \in \mathbb{N}$, let

$$\ell_i(n) = \min \left\{ p \in \mathbb{N} : \ p \geq (\gamma_1 + \cdots + \gamma_i) \frac{n}{\gamma_1} \right\},$$

and by convention set $\ell_0(n) = 0$. It is easy to check that in the ultrametric space $(X_1, d_{\vec{\gamma}})$, the closed ball centred at x of radius $e^{-\frac{n}{\gamma_1}}$ is given by

$$B(x, e^{-\frac{n}{\gamma_1}}) = \left\{ y \in X_1 : \ \Pi_i(y_{|\ell_i(n)}) = \Pi_i(x_{|\ell_i(n)}) \text{ for all } 1 \leq i \leq k \right\}.$$

Let $\mathcal{F}_n$ be the partition of X_1 into closed balls of radius $e^{-\frac{n}{\gamma_1}}$. For any finite Borel measure ν on X_1, the L^q-spectrum of ν can be defined as the concave mapping

$$\tau_\nu : \ q \in \mathbb{R} \mapsto \liminf_{n \to \infty} -\frac{\gamma_1}{n} \log \sum_{B \in \mathcal{F}_n} \nu(B)^q,$$

with the convention $0^q = 0$.

It is known that since $(X_1, d_{\vec{\gamma}})$ satisfies the Besicovich covering property, for ν-almost every $x \in X_1$, one has

$$\tau_\nu'(1^+) \leq \underline{\dim}_{\mathrm{loc}}(\nu, x) \leq \overline{\dim}_{\mathrm{loc}}(\nu, x) \leq \tau_\nu'(1^-),$$

so that the existence of $\tau_\nu'(1)$ implies the exact dimensionality of ν, with dimension equal to $\tau_\nu'(1)$ (see, e.g., [45]). Consequently, Theorem 2.2 follows from the following stronger result.

Theorem 3.1. *Suppose that $T(q) > -\infty$ for some $q > 1$. Conditional on $\mu \neq 0$, $\tau_\mu'(1)$ exists and equals $\dim_e^{\vec{\gamma}}(\mu)$, where $\dim_e^{\vec{\gamma}}(\mu)$ is defined as in (2.6).*

Recall that for $2 \leq i \leq k$, we defined ν_i as $\Pi_i \mathbb{E}(\mu) = \mathbb{E}(\Pi_i \mu)$. If ν is a Bernoulli product measure on X_i, we set

$$(3.1) \qquad \mathcal{T}_\nu(q) = -\log \sum_{b \in \mathcal{A}_i} \nu([b])^q \quad (q \geq 0).$$

Then Theorem 3.1 follows from the following proposition.

Proposition 3.2. *Suppose that $T(q) > -\infty$ for some $q > 1$. Let*

$$i_0 = \max\{2 \leq i \leq k : \ T'(1) \leq \mathcal{T}_{\nu_i}'(1)\},$$

with the convention $\max(\emptyset) = 1$. Then there exists $q_0 > 1$ and $c_0 \geq 0$ such that for all $q \in (0, q_0]$, we have

$$(3.2) \quad \mathbb{E}\left(\sum_{B \in \mathcal{F}_n} \mu(B)^q \right) = O\left(\exp\left(\ell_{i_0}(n)(c_0(q-1)^2 - T(q)) - \sum_{i=i_0+1}^{k} (\ell_i(n) - \ell_{i-1}(n))\mathcal{T}_{\nu_i}(q) \right) \right)$$

as $n \to \infty$, here we use the Landau big O notation. Moreover, c_0 can be taken equal to 0 if one restricts q to belong to $(0,1]$ or if $T'(1) \neq \mathcal{T}_{\nu_i}'(1)$ for all $2 \leq i \leq k$.

16

Assume that Proposition 3.2 holds. Then, a standard argument (see, e.g. [7, Lemma C]) yields that for any fixed $q \in (0, q_0]$, the following holds almost surely:

$$\limsup_{n \to \infty} \frac{\log \sum_{B \in \mathcal{F}_n} \mu(B)^q}{n} \leq \limsup_{n \to \infty} \frac{\ell_{i_0}(n)(c_0(q-1)^2 - T(q)) - \sum_{i=i_0+1}^{k}(\ell_i(n) - \ell_{i-1}(n))\mathcal{T}_{\nu_i}(q)}{n}$$

$$= \frac{\gamma_1 + \cdots + \gamma_{i_0}}{\gamma_1}(c_0(q-1)^2 - T(q)) - \sum_{i=i_0+1}^{k} \frac{\gamma_i}{\gamma_1} \mathcal{T}_{\nu_i}(q).$$

Then, by the convexity of the two sides as functions of q, the inequality holds almost surely for all $q \in (0, q_0]$. Multiplying both sides by $-\gamma_1$ yields, conditional on $\mu \neq 0$,

$$\tau_\mu(q) \geq -(\gamma_1 + \cdots + \gamma_{i_0})c_0(q-1)^2 + (\gamma_1 + \cdots + \gamma_{i_0})T(q) + \sum_{i=i_0+1}^{k} \gamma_i \mathcal{T}_{\nu_i}(q).$$

Since both sides of the above inequality are concave functions which coincide at $q = 1$ and the right hand side is differentiable at 1, we necessarily have that $\tau'_\mu(1)$ does exist and is equal to the derivative at 1 of the right hand side, namely

$$(\gamma_1 + \cdots + \gamma_{i_0})T'(1) + \sum_{i=i_0+1}^{k} \gamma_i \mathcal{T}'_{\nu_i}(1) = \dim_e^{\vec{\gamma}}(\mu).$$

Hence Proposition 3.2 implies Theorem 3.1.

In the remaining part of this section, we prove Proposition 3.2. We still need two additional lemmas.

Lemma 3.3. *Suppose that $T(q) > -\infty$ for some $q > 1$. Then, for all $q \in (1, 2)$ such that $T(q) > 0$, there exists a constant $C_q > 0$ such that for all $2 \leq i \leq k$, for all $n \geq 1$ one has*

$$\max\left(\mathbb{E} \sum_{U \in \mathcal{A}_i^n} \Pi_i \mu([U])^q, \ \mathbb{E} \sum_{U \in \mathcal{A}_i^n} \Pi_i \mu_n([U])^q\right) \leq C_q e^{-n \min(T(q), \mathcal{T}_{\nu_i}(q))}.$$

Proof. This is a direct consequence of [7, Corollary 5.2], in which the case $k = 2$ is considered. $\square$

Lemma 3.4. [49] *Let $(L_j)_{j \geq 1}$ be a sequence of centred independent real valued random variables. For every finite $I \subset \mathbb{N}$ and $q \in (1, 2]$,*

$$\mathbb{E}\left(\left|\sum_{i \in I} L_i\right|^q\right) \leq 2 \sum_{i \in I} \mathbb{E}(|L_i|^q).$$

Proof of Proposition 3.2. At first, note that the set of balls $\mathcal{F}_n$ is in bijection with the set $\prod_{i=1}^{k} \mathcal{A}_i^{\ell_i(n) - \ell_{i-1}(n)}$, since for any $x = (x_i)_{i=1}^{\infty} \in X_1$, if we set $U_i = \Pi_i(x_{\ell_{i-1}(n)+1} \cdots x_{\ell_i(n)})$, $1 \leq i \leq k$, and $U = (U_1, \ldots, U_k)$, then

$$B(x, e^{-\frac{n}{\gamma_1}}) = \left\{y \in X_1 : \Pi_i(T_1^{\ell_{i-1}(n)}y) \in [U_i] \text{ for all } 1 \leq i \leq k\right\}$$

$$= \bigcup_{(J_1, \ldots, J_k) \in \mathcal{J}_U} [J_1 \cdots J_k],$$

17

where

$$(3.3) \qquad \mathcal{J}_U := \left\{ (J_1, \ldots, J_k) \in \prod_{i=1}^{k} \mathcal{A}_1^{\ell_i(n) - \ell_{i-1}(n)} : \Pi_i(J_i) = U_i \text{ for all } 1 \leq i \leq k \right\}.$$

For $1 \leq i \leq k$, set $\mathcal{U}_n^{(i)} = \prod_{j=i}^{k} \mathcal{A}_j^{\ell_j(n) - \ell_{j-1}(n)}$. For $q \in \mathbb{R}_+$ we need to estimate from above the partition function

$$Z_{q,n} := \sum_{B \in \mathcal{F}_n} \mu(B)^q = \sum_{U \in \mathcal{U}_n^{(1)}} \left(\sum_{(J_1, J_2, \ldots, J_k) \in \mathcal{J}_U} \mu([J_1 \cdots J_k]) \right)^q.$$

For $1 \leq i \leq k$ and $U^{(i)} = (U_i, \cdots, U_k) \in \mathcal{U}_n^{(i)}$, set

$$\mathcal{J}_{U^{(i)}} = \left\{ (J_i, \ldots, J_k) \in \prod_{j=i}^{k} \mathcal{A}_1^{\ell_j(n) - \ell_{j-1}(n)} : \Pi_j(J_j) = U_j \text{ for all } i \leq j \leq k \right\}.$$

Also, set $\mathcal{U}_n^{(k+1)} = \{\epsilon\}$ and $\mathcal{J}_\epsilon = \{\epsilon\}$.

Then, for $1 \leq i \leq k$ and $(J_1, \ldots, J_i) \in \prod_{j=1}^{i} \mathcal{A}_1^{\ell_i(n) - \ell_{i-1}(n)}$, define the random variable

$$Z_{q,n}(J_1 \cdots J_i) = \sum_{U^{(i+1)} \in \mathcal{U}_n^{(i+1)}} \left(\sum_{(J_{i+1}, \ldots, J_k) \in \mathcal{J}_{U^{(i+1)}}} \mu^{J_1 \cdots J_i}([J_{i+1} \ldots J_k]) \right)^q.$$

Notice that $Z_{q,n}(\epsilon) = Z_{q,n}$ and $Z_{q,n}(J_1 \cdots J_k) = Y(J_1 \cdots J_k)^q$.

Due to the branching property associated with the measures μ^J, $J \in \mathcal{A}_1^*$, for all $0 \leq i \leq k-1$ we have

$$Z_{q,n}(J_1 \cdots J_i) = \sum_{U^{(i+1)} \in \mathcal{U}_n^{(i+1)}} \left(\sum_{\substack{J_{i+1} \in \mathcal{A}_1^{\ell_{i+1}(n) - \ell_i(n)} \\ \Pi_{i+1}(J_{i+1}) = U_{i+1}}} \mu_{\ell_{i+1}(n) - \ell_i(n)}^{J_1 \cdots J_i}([J_{i+1}]) S_{U^{(i+2)}}(J_1 \cdots J_i J_{i+1}) \right)^q,$$

where $U^{(i+2)} = (U_{i+2}, \ldots, U_k) \in \mathcal{U}_n^{(i+2)}$, and

$$S_{U^{(i+2)}}(J_1 \cdots J_i J_{i+1}) = \sum_{(J_{i+2}, \ldots, J_k) \in \mathcal{J}_{U^{(i+2)}}} \mu^{J_1 \cdots J_i J_{i+1}}([J_{i+2} \ldots J_k]).$$

Notice that the random variables $S_{U^{(i+2)}}(J_1 \cdots J_i J_{i+1})$, where $J_{i+1} \in \mathcal{A}_1^{\ell_{i+1}(n) - \ell_i(n)}$ and $\Pi_i(J_{i+1}) = U_{i+1}$, are independent and identically distributed, and independent of the σ-algebra generated by the $\mu_{\ell_{i+1}(n) - \ell_i(n)}^{J_1 \cdots J_i}(J_{i+1})$. Setting

$$L(J_{i+1}) = S_{U^{(i+2)}}(J_1 \cdots J_{i+1}) - \mathbb{E}(S_{U^{(i+2)}}),$$

where $\mathbb{E}(S_{U^{(i+2)}})$ stands for the common value of the $S_{U^{(i+2)}}(J_1 \cdots J_{i+1})$ expectations, we have, for $q > 1$:

$$
\mathbb{E}\left(\sum_{\substack{J_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)} \\ \Pi_{i+1}(J_{i+1})=U_{i+1}}} \mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1 \cdots J_i}([J_{i+1}]) \, S_{U^{(i+2)}}(J_1 \cdots J_i J_{i+1}) \right)^q
$$

$$
\leq 2^{q-1} \mathbb{E}\left(\sum_{\substack{J_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)} \\ \Pi_{i+1}(J_{i+1})=U_{i+1}}} \mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1 \cdots J_i}([J_{i+1}]) \right)^q \mathbb{E}(S_{U^{(i+2)}})^q
$$

$$
+ 2^{q-1} \mathbb{E}\left| \sum_{\substack{J_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)} \\ \Pi_{i+1}(J_{i+1})=U_{i+1}}} \mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1 \cdots J_i}([J_{i+1}]) \, L(J_{i+1}) \right|^q .
$$

Assuming that $q \in (1,2]$, we can apply Lemma 3.4 to the second term conditional on the σ-algebra generated by the $\mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1 \cdots J_i}(J_{i+1})$ and get

$$
\mathbb{E}\left| \sum_{\substack{J_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)} \\ \Pi_{i+1}(J_{i+1})=U_{i+1}}} \mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1 \cdots J_i}([J_{i+1}]) \, L(J_{i+1}) \right|^q
$$

$$
\leq 2^q \mathbb{E}\left(\sum_{\substack{J_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)} \\ \Pi_{i+1}(J_{i+1})=U_{i+1}}} \mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1 \cdots J_i}([J_{i+1}])^q \right) \mathbb{E}(|L|^q)
$$

$$
\leq 2^q \mathbb{E}\left(\sum_{\substack{J_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)} \\ \Pi_{i+1}(J_{i+1})=U_{i+1}}} \mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1 \cdots J_i}([J_{i+1}]) \right)^q \mathbb{E}(|L|^q) \quad \text{(using superadditivity)}
$$

$$
= 2^q \mathbb{E}\left(\Pi_{i+1} \mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1 \cdots J_i}([U_{i+1}])^q \right) \mathbb{E}(|L|^q),
$$

where $\mathbb{E}(|L|^q) = \mathbb{E}(|S_{U^{(i+2)}} - \mathbb{E}(S_{U^{(i+2)}})|^q) \leq 2^q \mathbb{E}(S_{U^{(i+2)}}^q)$, and $\mathbb{E}(S_{U^{(i+2)}}^q)$ is the common value of the $\mathbb{E}(S_{U^{(i+2)}}(J_1 \cdots J_{i+1})^q)$. Incorporating the last inequality in the previous one,

we get

$$\mathbb{E}\left(\sum_{\substack{J_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}\\\Pi_{i+1}(J_{i+1})=U_{i+1}}}\mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1\cdots J_i}(J_{i+1})\,S_{U^{(i+2)}}(J_1\cdots J_{i+1})\right)^q$$

$$\leq 2^{3q}\mathbb{E}\left(\Pi_{i+1}\mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1\cdots J_i}([U_{i+1}])^q\right)\mathbb{E}\left(S_{U^{(i+2)}}^q\right).$$

Then, taking an arbitrary element $\widetilde{J}_{i+1}$ in $\mathcal{A}_1^{\ell_{i+1}(n)-\ell_i(n)}$, we obtain

$$\mathbb{E}(Z_{q,n}(J_1\cdots J_i))$$

$$\leq 2^{3q}\sum_{U_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}}\mathbb{E}\big(\Pi_{i+1}\mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1\cdots J_i}([U_{i+1}])^q\big)\sum_{U^{(i+2)}\in\mathcal{U}_n^{(i+2)}}\mathbb{E}(S_{U^{(i+2)}}(J_1\cdots J_i\widetilde{J}_{i+1})^q)$$

$$=2^{3q}\sum_{U_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}}\mathbb{E}\big(\Pi_{i+1}\mu_{\ell_{i+1}(n)-\ell_i(n)}^{J_1\cdots J_i}([U_{i+1}])^q\big)\,\mathbb{E}(Z_{q,n}(J_1\cdots J_i\widetilde{J}_{i+1})).$$

Since $(\mu_p)_{p\geq 1}$ and $(\mu_p^{J_1\cdots J_i})_{p\geq 1}$ are identically distributed this yields

$$\mathbb{E}(Z_{q,n}(J_1\cdots J_i))$$

$$\leq 2^{3q}\sum_{U_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}}\mathbb{E}\big(\Pi_{i+1}\mu_{\ell_{i+1}(n)-\ell_i(n)}([U_{i+1}])^q\big)\,\mathbb{E}(Z_{q,n}(J_1\cdots J_i\widetilde{J}_{i+1})).$$

It follows that

$$\mathbb{E}(Z_{q,n})\leq 2^{3qk}\,\mathbb{E}(Y^q)\prod_{i=1}^{k}\mathbb{E}\left(\sum_{U_i\in\mathcal{A}_i^{\ell_i(n)-\ell_{i-1}(n)}}\Pi_i\mu_{\ell_i(n)-\ell_{i-1}(n)}([U_i])^q\right).$$

Let $q_1\in(1,2]$ such that $T(q)>0$ for all $q\in(1,q_1]$ (remember that $T(1)=0$ and $T'(1)>0$). Then, for all $q\in(1,q_1]$, the previous estimate combined with Lemma 3.3 yields

$$\mathbb{E}(Z_{q,n})\leq 2^{3qk}C_q^{k-1}\mathbb{E}(Y^q)\,\mathbb{E}\Big(\sum_{U_1\in\mathcal{A}_1^n}\mu_n([U_1])^q\Big)$$

$$\times\ \exp\Big(-\sum_{i=2}^{k}(\ell_i(n)-\ell_{i-1}(n))\min(T(q),\mathcal{T}_{\nu_i}(q))\Big)$$

$$=2^{3qk}C_q^{k-1}\mathbb{E}(Y^q)\exp\Big(-nT(q)-\sum_{i=2}^{k}(\ell_i(n)-\ell_{i-1}(n))\min(T(q),\mathcal{T}_{\nu_i}(q))\Big).$$

Finally, recall that $i_0=\max\{2\leq i\leq k:T'(1)\leq\mathcal{T}_{\nu_i}'(1)\}$ (note that for each i the number $\mathcal{T}_{\nu_i}'(1)$ is the measure-theoretic entropy of ν_i so that the sequence $(\mathcal{T}_{\nu_i}'(1))_{1\leq i\leq k}$ is non-increasing). Since T and the functions $\mathcal{T}_{\nu_i}$ are analytic near 1 and coincide at 1, for all $2\leq i\leq i_0$ there exist $q_{0,i}\in(1,q_1]$ and $c_i\geq 0$ such that for all $q\in(1,q_{0,i}]$ one has

$$\min(T(q),\mathcal{T}_{\nu_i}(q))\geq T(q)-c_i(q-1)^2,$$

with $c_i = 0$ if $T'(1) < \mathcal{T}'_{\nu_i}(1)$. Taking $c_0 = \max\{c_i : 2 \le i \le i_0\}$ and $q_0 = \min\{q_{0,i} : 2 \le i \le i_0\}$ yields (3.2).

Suppose now that $q \in (0, 1]$. We start with giving general estimate of $\mathbb{E}(Z_{q,n}(J_1 \cdots J_i))$. Using the subadditivity of $x \in \mathbb{R}_+ \mapsto x^q$ we have

$$Z_{q,n}(J_1 \cdots J_i) \le \sum_{\substack{U^{(i+1)} \in \mathcal{U}_n^{(i+1)}}} \sum_{\substack{J_{i+1} \in \mathcal{A}_1^{\ell_{i+1}(n) - \ell_i(n)} \\ \Pi_{i+1}(J_{i+1}) = U_{i+1}}} \mu_{\ell_{i+1}(n) - \ell_i(n)}^{J_1 \cdots J_i}([J_{i+1}])^q \, S_{U^{(i+2)}}(J_1 \cdots J_i J_{i+1})^q,$$

so

$$\mathbb{E}(Z_{q,n}(J_1 \cdots J_i)) \le \sum_{\substack{U_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n) - \ell_i(n)}}} \sum_{\substack{J_{i+1} \in \mathcal{A}_1^{\ell_{i+1}(n) - \ell_i(n)} \\ \Pi_{i+1}(J_{i+1}) = U_{i+1}}} \mathbb{E}(\mu_{\ell_{i+1}(n) - \ell_i(n)}^{J_1 \cdots J_i}(J_{i+1})^q)$$

$$\cdot \mathbb{E}\Big(\sum_{U^{(i+2)} \in \mathcal{U}_n^{(i+2)}} S_{U^{(i+2)}}(J_1 \cdots J_i \widetilde{J}_{i+1})^q \Big)$$

$$= \mathbb{E}\Big(\sum_{J_{i+1} \in \mathcal{A}_1^{\ell_{i+1}(n) - \ell_i(n)}} \mu_{\ell_{i+1}(n) - \ell_i(n)}^{J_1 \cdots J_i}(J_{i+1})^q \Big) \mathbb{E}(Z_{q,n}(J_1 \cdots J_i \widetilde{J}_{i+1}))$$

$$= \mathbb{E}\Big(\sum_{J_{i+1} \in \mathcal{A}_1^{\ell_{i+1}(n) - \ell_i(n)}} \mu_{\ell_{i+1}(n) - \ell_i(n)}(J_{i+1})^q \Big) \mathbb{E}(Z_{q,n}(J_1 \cdots J_i \widetilde{J}_{i+1}))$$

$$= \exp(-(\ell_{i+1}(n) - \ell_i(n))T(q)) \mathbb{E}(Z_{q,n}(J_1 \cdots J_i \widetilde{J}_{i+1})).$$

Starting from $\mathbb{E}(Z_{q,n}) = \mathbb{E}(Z_{q,n}(\epsilon))$ and iterating i_0 times the previous estimate we get

$$\mathbb{E}(Z_{q,n}) \le \Big(\prod_{i=1}^{i_0} \exp(-(\ell_i(n) - \ell_{i-1}(n))T(q)) \Big) \mathbb{E}(Z_{q,n}(\widetilde{J}_1 \cdots \widetilde{J}_{i_0}))$$

$$= \exp(-\ell_{i_0}(n)T(q)) \, \mathbb{E}(Z_{q,n}(\widetilde{J}_1 \cdots \widetilde{J}_{i_0})).$$

On the other hand, setting $\widetilde{J} = \widetilde{J}_1 \cdots \widetilde{J}_{i_0}$ and $\lambda(n) = \ell_{k-1}(n) - \ell_{i_0+1}(n)$, we can write

$$Z_{q,n}(\widetilde{J})$$

$$= \sum_{U^{(i_0+1)} \in \mathcal{U}_n^{(i_0+1)}} \left(\sum_{\substack{J' = J_{i_0+1} \cdots, J_{k-1} \\ \Pi_j(J_j) = U_j, \, \forall i_0+1 \le j \le k-1}} \mu_{\lambda(n)}^{\widetilde{J}}([J']) \sum_{J_k : \Pi_k(J_k) = U_k} \mu^{\widetilde{J}J'}([J_k]) \right)^q$$

$$= \sum_{U^{(i_0+1)} \in \mathcal{U}_n^{(i_0+1)}} \left(\sum_{\substack{J' = J_{i_0+1} \cdots J_{k-1} \\ \Pi_j(J_j) = U_j, \, \forall i_0+1 \le j \le k-1}} \mu_{\lambda(n)}^{\widetilde{J}}([J']) \, \nu_k([U_k]) \, X(\widetilde{J}J') \right)^q,$$

where

$$X(\widetilde{J}J') = \sum_{J_k : \Pi_k(J_k) = U_k} w^{\widetilde{J}J'}(J_k), \text{ with } w^{\widetilde{J}J'}(J_k) = \begin{cases} \dfrac{\mu^{\widetilde{J}J'}([J_k])}{\nu_k([U_k])} & \text{if } \nu_k([U_k]) > 0, \\ 0 & \text{otherwise} \end{cases}.$$

21

We can now use the independence of the random variables $X(\widetilde{J}J')$ with respect to the σ-algebra generated by the $\mu^{\widetilde{J}}_{\lambda(n)}([J'])$, conditioned with respect to this σ-algebra and use Jensen's inequality to get

$$\mathbb{E}(Z_{q,n}(\widetilde{J})) \leq \sum_{U^{(i_0+1)} \in \mathcal{U}_n^{(i_0+1)}} \mathbb{E}\left(\sum_{\substack{J'=J_{i_0+1},\ldots,J_{k-1} \\ \Pi_j(J_j)=U_j, \ \forall i_0+1 \leq j \leq k-1}} \mu^{\widetilde{J}}_{\lambda(n)}([J']) \, \nu_k([U_k]) \, \mathbb{E}(X(\widetilde{J}J')) \right)^q.$$

But by construction we have $\nu_k([U_k]) = \mathbb{E}\Big(\sum_{J_k: \, \Pi_k(J_k)=U_k} \mu^{\widetilde{J}J'}([J_k]) \Big)$, hence $\mathbb{E}(X(\widetilde{J}J')) = 1$. Setting

$$R = \sum_{(U_{i_0+1},\ldots,U_{k-1}) \in \prod_{i=i_0+1}^{k-1} \mathcal{A}_i^{\ell_i(n)-\ell_{i-1}(n)}} \mathbb{E}\left(\sum_{\substack{J'=J_{i_0+1}\cdots J_{k-1} \\ \Pi_j(J_j)=U_j, \ \forall i_0+1 \leq j \leq k-1}} \mu^{\widetilde{J}}_{\lambda(n)}([J']) \right)^q,$$

this yields

$$\mathbb{E}(Z_{q,n}(\widetilde{J})) \leq R \cdot \Big(\sum_{U_k \in \mathcal{A}_k^{\ell_k(n)-\ell_{k-1}(n)}} \nu_k([U_k])^q \Big) = R \cdot \exp(-(\ell_k(n) - \ell_{k-1}(n))\mathcal{T}_{\nu_k}(q)).$$

We can apply to R the same type of estimate as that for $\mathbb{E}(Z_{q,n}(\widetilde{J}))$, the only change being that $\mu^{\widetilde{J}J'}([J_k]) = \mu^{\widetilde{J}J_{i_0+1}\cdots J_{k-1}}([J_k])$ must be replaced by $\mu^{\widetilde{J}J_{i_0+1}\cdots J_{k-2}}_{\ell_{k-1}(n)-\ell_{k-2}(n)}(J_{k-1})$, and one now must use the fact that $\nu_{k-1} = \mathbb{E}\left(\Pi_{k-1} \mu^{\widetilde{J}J_{i_0+1}\cdots J_{k-2}}_{\ell_{k-1}(n)-\ell_{k-2}(n)} \right)$. Iterating we get

$$\mathbb{E}(Z_{q,n}(\widetilde{J})) \leq \exp\Big(- \sum_{i=i_0+1}^{k} (\ell_i(n) - \ell_{i-1}(n))\mathcal{T}_{\nu_i}(q) \Big),$$

and finally

$$\mathbb{E}(Z_{q,n}) \leq \exp\Big(- \ell_{i_0}(n)T(q) - \sum_{i=i_0+1}^{k} (\ell_i(n) - \ell_{i-1}(n))\mathcal{T}_{\nu_i}(q) \Big).$$

This completes the proof of the Proposition. $\qquad\square$

4. The Hausdorff dimension of K. Proof of Theorem 2.3

We have to optimise the weighted entropy $\dim_e^{\vec{\gamma}}(\mu)$ over the set of non-degenerate Mandelbrot measures μ supported on K; this will provide us with a sharp lower bound for $\dim_H K$. To do so, it is convenient to first relate $\dim_e(\mu)$ to $h_{\nu_i}(T_i)$ for all $1 \leq i \leq k$, where $\nu_i = \mathbb{E}(\Pi\mu)$. This is the purpose of Section 4.1. Then we identify at which point the maximum of weighted entropy dimension of Mandelbrot measures supported on K is reached. This constitutes Section 4.2. Section 4.3 quickly derives the sharp lower bound for $\dim_H K$. Finally, in Section 4.4 we develop a kind of variational principle to get the sharp upper bound for $\dim_H K$.

4.1. Mandelbrot measure as a kind of skewed product and decomposition of entropy dimension.

Let μ be a non-degenerate Mandelbrot measure jointly constructed with K and such that $K_\mu \subset K$ almost surely. As in Section 2.2, we denote by W the random vector used to generate μ. By construction, for any $1 \le i \le k$, the measure $\nu_i = \mathbb{E}(\Pi_i \mu)$ is the Bernoulli product measure on X_i associated with the probability vector $p^{(i)} = (p_b^{(i)})_{b \in \mathcal{A}_i}$, where

$$p_b^{(i)} = \sum_{a \in \mathcal{A}_1 \colon [a] \subset \Pi_i^{-1}([b])} \mathbb{E}(W_a),$$

and one has $\nu_i = \pi_{i-1}\nu_{i-1}$ for $i \ge 2$. We also define, for $b \in \mathcal{A}_i$,

$$V_b^{(i)} = (V_{b,a}^{(i)})_{a \in \mathcal{A}_1 \colon [a] \subset \Pi_i^{-1}([b])} = \begin{cases} \nu_i([b])^{-1} (W_a)_{a \in \mathcal{A}_1 \colon [a] \subset \Pi_i^{-1}([b])} & \text{if } \nu_i([b]) > 0 \\ 0 & \text{otherwise} \end{cases},$$

so that for all $a \in \mathcal{A}_1$, for all $1 \le i \le k$, we have the multiplicative decomposition

$$W_a = \nu_i(\Pi_i[a]) \cdot V_{\Pi_i(a),a}^{(i)}.$$

For $1 \le i \le k$ and $b \in \mathcal{A}_i$, set

$$(4.1) \qquad T_{V_b^{(i)}}(q) = -\log \mathbb{E}\left(\sum_{a \in \mathcal{A}_1 \colon [a] \subset \Pi_i^{-1}([b])} (V_{b,a}^{(i)})^q \right) \quad (q \ge 0),$$

with the conventions $0^0 = 0$ and $\log(0) = -\infty$. In particular

$$(4.2) \qquad T_{V_b^{(i)}}(0) = -\log \mathbb{E}(N_b^{(i)}(W)),$$

where

$$N_b^{(i)}(W) = \#\{a \in \mathcal{A}_1 \colon [a] \subset \Pi_i^{-1}([b]), \ W_a > 0\}.$$

One can check that

$$(4.3) \qquad e^{-T_W(q)} = \sum_{\substack{b \in \mathcal{A}_i \\ \nu_i([b]) > 0}} \nu_i([b])^q e^{-T_{V_b^{(i)}}(q)},$$

from what it follows, after differentiating at 1, that

$$(4.4) \qquad \dim_e(\mu) = h_{\nu_i}(T_i) + \dim_e(\mu | \nu_i),$$

where

$$h_{\nu_i}(T_i) = -\sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \nu_i([b])$$

is the entropy of the T_i-invariant measure ν_i and

$$\dim_e(\mu | \nu_i) := \sum_{\substack{b \in \mathcal{A}_i \\ \nu_i([b]) > 0}} \nu_i([b]) T'_{V_b^{(i)}}(1) = \sum_{\substack{b \in \mathcal{A}_i \\ \nu_i([b]) > 0}} \nu_i([b]) \left(- \sum_{a \in \mathcal{A}_1 \colon [a] \subset \Pi_i^{-1}([b])} \mathbb{E}(V_{b,a}^{(i)} \log V_{b,a}^{(i)}) \right)$$

must be thought of as the relative entropy dimension of μ given ν_i whenever this number is non-negative.

Among the Mandelbrot measures supported on K, special ones will play a prominent role. We introduce them now.

Recall that for $1 \leq i \leq k$ and $b \in \mathcal{A}_i = \Pi_i(\mathcal{A}_1)$, we defined

$$N_b^{(i)} = \#\{a \in \mathcal{A}_1 : [a] \subset \Pi_i^{-1}([b]), [a] \cap K \neq \emptyset\}$$

and we also defined the set $\widetilde{\mathcal{A}}_i = \{b \in \mathcal{A}_i : \mathbb{E}(N_b^{(i)}) > 0\}$.

For $a \in \mathcal{A}_1$ such that $[a] \subset \Pi_i^{-1}([b])$ let

$$\widetilde{V}_{b,a}^{(i)} = \begin{cases} (\mathbb{E}(N_b^{(i)}))^{-1} & \text{if } b \in \widetilde{\mathcal{A}}_i \text{ and } [a] \cap K \neq \emptyset \\ 0 & \text{otherwise} \end{cases}.$$

If ν_i is a Bernoulli product measure on X_i, and W_a is taken equal to $\widetilde{W}_a = \nu_i([b])\widetilde{V}_{b,a}^{(i)}$ for all $a \in \mathcal{A}_1$ such that $[a] \subset \Pi_i^{-1}([b])$, and if $T'_{\widetilde{W}}(1) > 0$, we get a new Mandelbrot measure that we denote by μ_{ν_i}. By construction, $\nu_i = \mathbb{E}(\Pi_i \mu_{\nu_i})$, and

$$(4.5) \qquad \dim_e(\mu_{\nu_i}|\nu_i) = \sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \mathbb{E}(N_b^{(i)}).$$

The following remarks point out some properties which will play key roles in our proof of Theorem 2.3.

Remark 4.1. *Given a non-degenerate Mandelbrot measure μ supported on K and $1 \leq i \leq k$, for each $b \in \mathcal{A}_i$ such that $\nu_i([b]) > 0$, the function $T_{V_b^{(i)}}$ is concave, takes value 0 at 1 and due to (4.2) one has $T_{V_b^{(i)}}(0) = -\log \mathbb{E}(N_b^{(i)}(W)) \geq -\log \mathbb{E}(N_b^{(i)})$, so $T_{V_b^{(i)}}$ is bounded from below by the linear function $T_{\widetilde{V}_b^{(i)}} : q \mapsto (q-1)\log \mathbb{E}(N_b^{(i)})$. Consequently, $T'_{V_b^{(i)}}(1) \leq T'_{\widetilde{V}_b^{(i)}}(1) = \log \mathbb{E}(N_b^{(i)})$. It then follows from (4.4) that*

$$\begin{aligned} T'_{\widetilde{W}}(1) &= h_{\nu_i}(T_i) + \sum_{b \in \mathcal{A}_i : \nu_i([b])>0} \nu_i([b]) T'_{\widetilde{V}_b^{(i)}}(1) \\ &\geq h_{\nu_i}(T_i) + \sum_{b \in \mathcal{A}_i : \nu_i([b])>0} \nu_i([b]) T'_{V_b^{(i)}}(1) \\ &= T'_W(1), \end{aligned}$$

with equality only if $T_{V_b^{(i)}}$ equals $T_{\widetilde{V}_b^{(i)}}$ whenever $\nu_i([b]) > 0$. In particular, $T_{V_b^{(i)}}$ must be affine. Remembering (4.1), and differentiating twice shows, after an application of Cauchy-Schwarz inequality, that this implies that $\widetilde{V}_{b,a}^{(i)}$ equals $\mathbf{1}_{\mathbb{R}_+^*}(W_a)/\mathbb{E}(N_b^{(i)}(W))$ almost surely. Consequently, $T_{V_b^{(i)}}$ equals $T_{\widetilde{V}_b^{(i)}}$ if and only if $V_b^{(i)} = \widetilde{V}_b^{(i)}$ almost surely, for all b such that $\nu_i([b]) > 0$, i.e. $W = \widetilde{W}$ almost surely. In particular, either $T'_{\widetilde{W}}(1) = T'_W(1)$ and μ_{ν_i} is non-degenerate since $\mu_{\nu_i} = \mu$ almost surely, or $T'_{\widetilde{W}}(1) > T'_W(1) \geq 0$ so that μ_{ν_i} is non-degenerate (recall that since μ is non-degenerate, $T'_W(1) = 0$ only in the case that μ is a Dirac measure).

Recall that for $i = 1, \ldots, k$, $X_i = \mathcal{A}_i^{\mathbb{N}}$ and $\widetilde{X}_i = (\widetilde{\mathcal{A}}_i)^{\mathbb{N}}$, where

$$\widetilde{\mathcal{A}}_i := \{b \in \mathcal{A}_i : \mathbb{E}(N_b^{(i)}) > 0\}.$$

Remark 4.2. *The reader will also check that when μ_{ν_i} is non-degenerate, for all $1 \le i' \le i-1$, denoting $\mathbb{E}(\Pi_{i'}\mu_{\nu_i})$ by $\nu_{i'}$, one also has $\widetilde{W}_a = \nu_{i'}(b')\widetilde{V}_{b',a}^{(i')}$ for all $b' \in \mathcal{A}_{i'}$ and $a \in \mathcal{A}_1$ such that $[a] \subset \Pi_{i'}^{-1}([b])$. Consequently, $\mu_{\nu_{i'}} = \mu_{\nu_i}$ and*

$$\dim_e(\mu_{\nu_i}) = h_{\nu_i}(T_i) + \sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \mathbb{E} N_b^{(i)} = h_{\nu_{i'}}(T_{i'}) + \sum_{b' \in \mathcal{A}_{i'}} \nu_{i'}([b']) \log \mathbb{E}(N_{b'}^{(i')})$$

(with the convention $0 \times (-\infty) = 0$). Also, if ν_i is fully supported on $\widetilde{X}_i$, then $K_{\mu_{\nu_i}} = K$ almost surely since by construction a component $\widetilde{W}_a$ of the random vector $\widetilde{W}$ associated with μ_{ν_i} as in Section 4.1 vanishes exactly when $a \notin A(\omega)$.

Definition 4.3. *We denote by $\mathcal{B}_i$ the set of Bernoulli product measures on X_i which are supported on $\widetilde{X}_i$. Also, if ρ_i is a Bernoulli product measure defined on $\widetilde{X}_i$, we also denote by ρ_i the Bernoulli product measure on X_i defined by assigning the mass $\rho_i(B \cap \widetilde{X}_i)$ to any Borel subset B of X_i.*

Remark 4.4. *Another important fact, which follows from the previous remarks, is that given $1 \le i \le k$ and $\rho_i \in \mathcal{B}_i$ such that $h_{\rho_i}(T_i) + \sum_{b \in \widetilde{\mathcal{A}}_i} \rho_i([b]) \log \mathbb{E} N_b^{(i)} > 0$, then μ_{ρ_i} is the unique Mandelbrot measure μ jointly defined with K and such that if $\rho = \mathbb{E}(\mu)$, then $h_\rho(T_1) + \sum_{b \in \widetilde{\mathcal{A}}_1} \rho([b]) \log \mathbb{E} N_b^{(1)} = h_{\rho_i}(T_i) + \sum_{b \in \mathcal{A}_i} \rho_i([b]) \log \mathbb{E} N_b^{(i)}$.*

On the other hand, it also holds that if $\rho_i \in \mathcal{B}_i$, then

$$\sup \left\{ h_\rho(T_1) + \sum_{b \in \widetilde{\mathcal{A}}_1} \rho([b]) \log \mathbb{E} N_b^{(1)} : \rho \in \mathcal{B}_1, \ \Pi_i\rho = \rho_i \right\} = h_{\rho_i}(T_i) + \sum_{b \in \mathcal{A}_i} \rho_i([b]) \log \mathbb{E} N_b^{(i)},$$

and the supremum is uniquely attained. See e.g. [42] for a proof (or use [12, Lemma 1.1] for a direct proof).

4.2. An optimisation problem.

The following result invokes notions introduced to state Theorem 2.3 and in the previous Section 4.1. Recall Definition 4.3.

Theorem 4.5. *Let*

$$M^{\vec{\gamma}} = \max\{\dim_e^{\vec{\gamma}}(\mu) : \mu \text{ is a positive Mandelbrot measure supported on } K\}.$$

One has

$$M^{\vec{\gamma}} = \begin{cases} P_{i_0}(\theta_{i_0}) & \text{if } i_0 \le k \\ P_k(1) & \text{if } i_0 = k+1 \end{cases},$$

and the maximum is uniquely attained at $\mu_{\nu_{i_0}}$, where ν_{i_0} is the equilibrium state of P_{i_0} at θ_0 and $\mu_{\nu_{i_0}}$ is the Mandelbrot measure associated to ν_{i_0} as in Section 4.1.

Before proving Theorem 4.5, we first recall a version of von Neumann's min-max theorem and establish two useful corollaries of it.

4.2.1. A generalised version of von Neumann's min-max theorem, and some applications.

In this part, we state a generalised version of von Neumann's min-max theorem, which was proved by Ky Fan [24].

We first introduce some definitions. Let f be a real function defined on the product set $X \times Y$ of two arbitrary sets X, Y (not necessarily topologized). f is said to be convex

25

on X, if for any two elements x_1, $x_2 \in X$ and $t \in [0,1]$, there exists an element $x_0 \in X$ such that $f(x_0, y) \leq tf(x_1, y) + (1-t)f(x_2, y)$ for all $y \in Y$. Similarly f is said to be concave on Y, if for any two elements y_1, $y_2 \in X$ and $t \in [0,1]$, there exists an element $y_0 \in Y$ such that $f(x, y_0) \geq tf(x, y_1) + (1-t)f(x, y_2)$ for all $x \in X$.

Theorem 4.6. [24, Theorem 2]. *Let X be a compact Hausdorff space and Y an arbitrary set (not topologized). Let f be a real-valued function on $X \times Y$ such that for every $y \in Y$, $f(x, y)$ is lower semi-continuous on X. If f is convex on X and concave on Y, then*

$$(4.6) \qquad \min_{x \in X} \sup_{y \in Y} f(x, y) = \sup_{y \in Y} \min_{x \in X} f(x, y).$$

Now we give a specific application of Theorem 4.6. Let Y be a compact convex subset of a topological vector space. Let $\ell \in \mathbb{N}^*$ and $g_1, \ldots, g_\ell$ be real concave functions defined on Y. Set

$$\Delta_\ell = \left\{ (q_1, \ldots, q_\ell) \in \mathbb{R}^\ell : \ q_i \geq 0 \text{ and } \sum_{i=1}^\ell q_i = 1 \right\}.$$

Define $P : \Delta_\ell \to \mathbb{R}$ by

$$P(q_1, \ldots, q_\ell) = \sup \left\{ \sum_{i=1}^\ell q_i g_i(y) : \ y \in Y \right\}.$$

Clearly P is a convex function on Δ_ℓ. As a consequence of Theorem 4.6 we have the following.

Corollary 4.7. *Under the above setting, we have*

$$(4.7) \qquad \min_{(q_1, \ldots, q_\ell) \in \Delta_\ell} P(q_1, \ldots, q_\ell) = \sup_{y \in Y} \min \{g_1(y), \ldots, g_\ell(y)\}.$$

Proof. Define $f : \Delta_\ell \times Y \to \mathbb{R}$ by

$$f(\mathbf{q}, y) = \sum_{i=1}^\ell q_i g_i(y) \quad \text{for } \mathbf{q} = (q_1, \ldots, q_\ell) \in \Delta_\ell, \ y \in Y.$$

Clearly, f is convex and continuous in $\mathbf{q} = (q_1, \ldots, q_\ell)$ and concave in y. Notice that for every $y \in Y$,

$$\min \{g_1(y), \ldots, g_\ell(y)\} = \min_{(q_1, \ldots, q_\ell) \in \Delta_\ell} \sum_{i=1}^\ell q_i g_i(y) = \min_{(q_1, \ldots, q_\ell) \in \Delta_\ell} f((q_1, \ldots, q_\ell), y).$$

Applying Theorem 1.1 to f (in which we take $X = \Delta_\ell$) yields (4.7). $\qquad \square$

Next we rewrite (4.7) as

$$\min_{\mathbf{q} \in \Delta_\ell} \sup_{y \in Y} \mathbf{q} \cdot \mathbf{g}(y) = \sup_{y \in Y} \min_{\mathbf{q} \in \Delta_\ell} \mathbf{q} \cdot \mathbf{g}(y) = \sup_{y \in Y} \min \{g_1(y), \ldots, g_\ell(y)\},$$

where $\mathbf{q} = (q_1, \ldots, q_\ell)$ and $\mathbf{g}(y) := (g_1(y), \ldots, g_\ell(y))$.

Corollary 4.8. *Assume in addition that $g_1, \ldots, g_\ell$ are strictly concave and continuous (or upper semi-continuous) on Y. Then the following properties hold:*

(i) *For any* $\mathbf{q} \in \Delta_\ell$, *there is a unique* $y = y_\mathbf{q} \in Y$ *which attains the supremum in the variational principle*

(4.8) $$P(\mathbf{q}) = \sup_{y \in Y} \mathbf{q} \cdot \mathbf{g}(y).$$

(ii) *Write* $L = \min_{\mathbf{q} \in \Delta_\ell} P(\mathbf{q})$. *Then there is a unique* $y \in Y$ *such that*

(4.9) $$\min\{g_1(y), \ldots, g_\ell(y)\} = L.$$

Moreover if P *takes this minimum at a point* $\mathbf{q}^* = (q_1^*, \ldots, q_\ell^*) \in \Delta_\ell$, *then* $y_{\mathbf{q}^*} = y$ *and furthermore,* $q_j^* = 0$ *if* $g_j(y) > L$.

Proof. Part (i) follows from the assumptions that g_i are upper semi-continuous and strictly concave. Specifically, the first assumption ensures that the supremum in (4.8) is attainable. The second one ensures that it is uniquely attained.

To see (ii), first notice that the uniqueness of y follows from the fact that L equals $\sup_{y \in Y} \min\{g_1(y), \ldots, g_\ell(y)\}$, and from the strict concavity of $\min\{g_1, \ldots, g_\ell\}$, which follows from the strict concavity of g_i. To prove the remaining statements of part (ii), suppose that $L = P(\mathbf{q}^*)$ for some $\mathbf{q}^* = (q_1^*, \ldots, q_\ell^*) \in \Delta_\ell$. By (i), there is a unique $y_{\mathbf{q}^*} \in Y$ so that $P(\mathbf{q}^*) = \mathbf{q}^* \cdot \mathbf{g}(y_{\mathbf{q}^*})$. Recall that $y \in Y$ is the unique element so that (4.9) holds. Clearly $\mathbf{q}^* \cdot \mathbf{g}(y) \geq L$. We prove that $y_{\mathbf{q}^*} = y$ by contradiction. If $y \neq y_{\mathbf{q}^*}$, then $g_i((y + y_{\mathbf{q}^*})/2) > (g_i(y) + g_i(y_{\mathbf{q}^*}))/2$ for all i so

$$P(\mathbf{q}^*) \geq \mathbf{q}^* \cdot \mathbf{g}((y + y_{\mathbf{q}^*})/2) > (\mathbf{q}^* \cdot \mathbf{g}(y) + \mathbf{q}^* \cdot \mathbf{g}(y_{\mathbf{q}^*}))/2 \geq L,$$

leading to a contradiction. Hence $y = y_{\mathbf{q}^*}$. Finally, since

$$\mathbf{q}^* \cdot \mathbf{g}(y) = L = \min\{g_1(y), \ldots, g_\ell(y)\},$$

it follows that $q_j^* = 0$ if $g_j(y) > L$. $\qquad\square$

4.2.2. **Proof of Theorem 4.5.** For $1 \leq i \leq k$, we identify $\mathcal{B}_i$, the set of Bernoulli product measures on X_i which are supported on $\widetilde{X}_i$, with $\mathcal{P}(\widetilde{A}_i)$, the compact convex set of probability measures on $\widetilde{A}_i$. An element $\nu_i \in \mathcal{B}_i$ is said to be *fully supported on* $\widetilde{A}_i$ if $\nu_i(b) \neq 0$ for every $b \in \widetilde{A}_i$. The projection π_i originally defined from $\mathcal{A}_i$ to $\mathcal{A}_{i+1}$ maps $\widetilde{A}_i$ to $\widetilde{A}_{i+1}$ ($i = 1, \ldots, k-1$). For $\nu_i \in \mathcal{B}_i$ let $H(\nu_i)$ denote the Shannon entropy of ν_i, which coincides with $h_{\nu_i}(T_i)$ but the notation $H(\nu_i)$ will be lighter.

Let $\widetilde{\phi} : \widetilde{A}_1 \to \mathbb{R}$ be defined by

$$\widetilde{\phi}(a) = \log \mathbb{E}(N_a^{(1)}), \quad q \in \widetilde{A}_1.$$

We identify $\widetilde{\phi}$ with the potential defined on $\widetilde{X}_1$ by $x \mapsto \widetilde{\phi}(x_1)$.

Write $\mathcal{B}_1^* := \{\nu \in \mathcal{B}_1 : H(\nu) + \nu(\widetilde{\phi}) > 0\}$. The proof of Theorem 4.5 is based on Remarks 4.1 to 4.4, Corollary 4.8 as well as the following two facts:

Lemma 4.9.

$$M^{\vec{\gamma}} = \max\left\{\gamma_1(H(\nu) + \nu(\widetilde{\phi})) + \sum_{i=2}^{k} \gamma_i \min\{H(\nu) + \nu(\widetilde{\phi}), H(\Pi_i\nu)\} : \nu \in \mathcal{B}_1^*\right\}$$

(4.10) $$= \max\left\{\gamma_1(H(\nu) + \nu(\widetilde{\phi})) + \sum_{i=2}^{k} \gamma_i \min\{H(\nu) + \nu(\widetilde{\phi}), H(\Pi_i\nu)\} : \nu \in \mathcal{B}_1\right\}.$$

Proof. At first, note that for any non-degenerate non-atomic Mandelbrot measure μ, $\dim_e(\mu) = H(\nu) + \nu(\widetilde{\phi}) > 0$, where $\nu = \mathbb{E}(\mu)$. Then, notice that if $\nu \in \mathcal{B}_1 \backslash \mathcal{B}_1^*$, we have $H(\nu) + \nu(\widetilde{\phi}) \leq 0$ and so

$$\gamma_1(H(\nu) + \nu(\widetilde{\phi})) + \sum_{i=2}^{k} \gamma_i \min\{H(\nu) + \nu(\widetilde{\phi}), H(\Pi_i \nu)\} \leq 0.$$

However we know that $M^{\vec{\gamma}} > 0$ (which follows from our assumption $\mathbb{E}(\#A) > 1$ and by considering the branching measure, i.e. the Mandelbrot measure on K associated with the random vector $W = (\mathbf{1}_A(a)/\mathbb{E}(\#A))_{a \in \mathcal{A}_1})$, so the maximum must be taken over $\mathcal{B}_1^*$. $\square$

Now we are ready to prove Theorem 4.5.

Proof of Theorem 4.5. For $i = 1, \ldots, k$, define $g_i : \mathcal{B}_1 \to \mathbb{R}$ by

$$(4.11) \qquad g_i(\nu) = (\gamma_1 + \ldots + \gamma_i)(H(\nu) + \nu(\widetilde{\phi})) + \sum_{j=i+1}^{k} \gamma_j H(\Pi_j \nu).$$

Then g_i are strictly concave functions on $\mathcal{B}_1$ (identified with $\mathcal{P}(\widetilde{A}_1)$). Clearly for each $\nu \in \mathcal{B}_1$,

$$\gamma_1(H(\nu) + \nu(\widetilde{\phi})) + \sum_{i=2}^{k} \gamma_i \min\{H(\nu) + \nu(\widetilde{\phi}), H(\Pi_i \nu)\} = \min_{1 \leq i \leq k} g_i(\nu).$$

Hence by (4.10), we obtain

$$(4.12) \qquad M^{\vec{\gamma}} = \sup_{\nu \in \mathcal{B}_1} \min\{g_1(\nu), \ldots, g_k(\nu)\}.$$

Define $P : \Delta_k \to \mathbb{R}$ by

$$(4.13) \qquad P(q_1, \ldots, q_k) = \sup_{\nu \in \mathcal{B}_1} \sum_{i=1}^{k} q_i g_i(\nu).$$

Applying Corollaries 4.7-4.8, we see that

$$M^{\vec{\gamma}} = \min_{(q_1, \ldots, q_k) \in \Delta_k} P(q_1, \ldots, q_k)$$

and that the supremum in (4.12) is uniquely attained at some element of $\mathcal{B}_1$, say ν; moreover, ν is the unique equilibrium state for $P(\mathbf{q}^*)$, provided that P takes its minimum at $\mathbf{q}^* = (q_1^*, \ldots, q_k^*)$.

Fix such a $\mathbf{q}^*$. Since ν is the equilibrium state for $P(\mathbf{q}^*)$, it must be fully supported on $\widetilde{A}_1$ (see [6]). By the strong concavity of the entropy function H, it follows that

$$(4.14) \qquad H(\Pi_i \nu) \geq H(\Pi_{i+1} \nu) \quad \text{for } i = 1, \ldots, k-1;$$

and the equality holds only if $\pi_i : \widetilde{\mathcal{A}}_i \to \widetilde{\mathcal{A}}_{i+1}$ is injective. So either one of the following 3 cases occurs: (i) $H(\nu) + \nu(\widetilde{\phi}) \geq H(\Pi_2 \nu)$; (ii) there exists a unique $i \in \{2, \ldots, k-1\}$ so that $H(\Pi_{i+1} \nu) \leq H(\nu) + \nu(\widetilde{\phi}) < H(\Pi_i \nu)$; (iii) $H(\nu) + \nu(\widetilde{\phi}) < H(\Pi_k \nu)$.

If Case (i) occurs, then $g_1(\nu) \leq g_2(\nu) \leq g_3(\nu) \leq \cdots \leq g_k(\nu)$, and by Corollary 4.8(ii), we have $q_j^* = \ldots = q_k^* = 0$ where j is the smallest index i such that $g_i(\nu) > g_2(\nu)$, if any;

if there is no such an index, set $j = k + 1$. So if $j = 3$, then $P(\mathbf{q}^*) = P(q_1^*, q_2^*, 0, \ldots, 0)$. Below we assume that $j > 3$. Since $g_2(\nu) = \ldots = g_{j-1}(\nu)$, it follows that

$$
(4.15) \qquad\qquad H(\Pi_2 \nu) = H(\Pi_3 \nu) = \ldots = H(\Pi_{j-1} \nu),
$$

and

$$
(4.16) \qquad\qquad P(\mathbf{q}^*) = P(q_1^*, q_2^*, \ldots, q_{j-1}^*, 0 \ldots, 0).
$$

By (4.15) we see that the mappings $\pi_2 : \widetilde{\mathcal{A}}_2 \to \widetilde{\mathcal{A}}_3, \ldots, \pi_{j-2} : \widetilde{\mathcal{A}}_{j-2} \to \widetilde{\mathcal{A}}_{j-1}$ are injective, which implies that

$$
(4.17) \qquad\qquad H(\Pi_2 \mu) = H(\Pi_3 \mu) = \cdots = H(\Pi_{j-1} \mu) \quad \text{for all } \mu \in \mathcal{B}_1.
$$

Now $P(q_1^*, \cdots, q_{j-1}^*, 0, \ldots, 0) = \sup_{\mu \in \mathcal{B}_1} (q_1^* g_1(\mu) + \cdots + q_{j-1}^* g_{j-1}(\mu))$. According to (4.17), we have

$$
g_1(\mu) = \gamma_1(H(\mu) + \mu(\widetilde{\phi})) + (\gamma_2 + \cdots + \gamma_{j-1})H(\Pi_2 \mu) + \sum_{p=j}^{k} \gamma_p H(\Pi_p \mu),
$$

$$
g_2(\mu) = (\gamma_1 + \gamma_2)(H(\mu) + \mu(\widetilde{\phi})) + (\gamma_3 + \cdots + \gamma_{j-1})H(\Pi_2 \mu) + \sum_{p=j}^{k} \gamma_p H(\Pi_p \mu),
$$

$$
\cdots
$$

$$
g_{j-1}(\mu) = (\gamma_1 + \cdots + \gamma_{j-1})(H(\mu) + \mu(\widetilde{\phi})) + \sum_{p=j}^{k} \gamma_p H(\Pi_p \mu).
$$

It follows that for $(q_1, \ldots, q_{j-1}) \in \Delta_{j-1}$,

$$
\sup_{\mu \in \mathcal{B}_1} \sum_{i=1}^{j-1} q_i g_i(\mu) = \sup_{\mu \in \mathcal{B}_1} \sum_{t=1}^{j-1} q_t(\gamma_1 + \cdots + \gamma_t)(H(\mu) + \mu(\widetilde{\phi})) + \sum_{s=1}^{j-2} q_s(\gamma_{s+1} + \cdots + \gamma_{j-1})H(\Pi_2 \mu)
$$

$$
+ \sum_{p=j}^{k} \gamma_p H(\Pi_p \mu)
$$

$$
= \sup_{\eta \in \mathcal{B}_2} \sum_{t=1}^{j-1} q_t(\gamma_1 + \cdots + \gamma_t)(H(\eta) + \eta(\widetilde{\phi}_2)) + \sum_{s=1}^{j-2} q_s(\gamma_{s+1} + \cdots + \gamma_{j-1})H(\eta)
$$

$$
+ \sum_{p=j}^{k} \gamma_p H(\Pi_{2,p} \eta) \qquad (\text{where } \Pi_{2,p} := \pi_{p-1} \circ \cdots \circ \pi_2)
$$

$$
= \sup_{\eta \in \mathcal{B}_2} \sum_{t=1}^{j-1} q_t(\gamma_1 + \cdots + \gamma_t)\eta(\widetilde{\phi}_2) + (\gamma_1 + \cdots + \gamma_{j-1})H(\eta) + \sum_{p=j}^{k} \gamma_p H(\Pi_{2,p} \eta),
$$

where in the second equality, we used the fact that

$$
\sup\{H(\mu) + \mu(\widetilde{\phi}) : \mu \in \mathcal{B}_1, \Pi_2 \mu = \eta\} = H(\eta) + \eta(\widetilde{\phi}_2)
$$

for each $\eta \in \mathcal{B}_2$ (see Remark 4.4). Hence $P(q_1, \ldots, q_{j-1}, 0, \ldots, 0)$ depends only on $\sum_{t=1}^{j-1} q_t(\gamma_1 + \cdots + \gamma_t)$. Clearly, there exists $2 < j' \leq j$ such that

$$
\gamma_1 + \cdots + \gamma_{j'-2} \leq \sum_{t=1}^{j-1} q_t(\gamma_1 + \cdots + \gamma_t) \leq \gamma_1 + \cdots + \gamma_{j'-1}
$$

29

so there exists $(p_{j'-2}, p_{j'-1}) \in \Delta_2$ such that

$$\sum_{t=1}^{j-1} q_t(\gamma_1 + \cdots + \gamma_t) = p_{j'-2}(\gamma_1 + \cdots + \gamma_{j'-2}) + p_{j'-1}(\gamma_1 + \cdots + \gamma_{j'-1})$$

and so

$$P(q_1, \ldots, q_{j-1}, 0, \ldots, 0) = P(\underbrace{0, \ldots, 0}_{(j'-3) \text{ times}}, p_{j'-2}, p_{j'-1}, 0, \ldots, 0).$$

Applying this fact to the case when $(q_1, \ldots, q_{j-1}) = (q_1^*, \ldots, q_{j-1}^*)$, we see that there exist $2 < j' \leq j$ and $(p_{j'-2}, p_{j'-1}) \in \Delta_2$ so that

$$P(\mathbf{q}^*) = P(q_1^*, \ldots, q_{j-1}^*, 0, \ldots, 0) = P(\underbrace{0, \ldots, 0}_{(j'-3) \text{ times}}, p_{j'-2}, p_{j'-1}, 0, \ldots, 0).$$

If Case (ii) occurs, then either (a) $H(\Pi_{i+1}\nu) < H(\nu) + \nu(\widetilde{\phi}) < H(\Pi_i \nu)$, or (b) $H(\Pi_{i+p+1}\nu) < H(\Pi_{i+p}\nu) = \ldots = H(\Pi_{i+1}\nu) = H(\nu) + \nu(\widetilde{\phi}) < H(\Pi_i \nu)$ for some $p \geq 1$. In the subcase (a), $g_1(\nu) > \ldots > g_i(\nu) < g_{i+1}(\nu) \leq \ldots \leq g_k(\nu)$, so by Corollary 4.8(ii), $q_j^* = 0$ for $j \neq i$, i.e. $\mathbf{q}^* = (0, \ldots, 0, q_i^*, 0, \ldots, 0)$ with $q_i^* = 1$. In the subcase (b), we have

$$q_1(\nu) > \ldots > g_i(\nu) = g_{i+1}(\nu) = \cdots = g_{i+p}(\nu) < g_{i+p+1}(\nu) \leq \ldots \leq g_k(\nu),$$

so by Corollary 4.8(ii), $q_j^* = 0$ if $j < i$ or $j > i + p$, i.e.

$$P(\mathbf{q}^*) = P(\underbrace{0, \ldots, 0}_{(i-1) \text{ times}}, q_i^*, \ldots, q_{i+p}^*, 0, \ldots, 0).$$

Following an argument similar as in the first case, we can show that there exist $1 \leq p' \leq p$ and $(q_{i+p'-1}, q_{i+p'}) \in \Delta_2$ so that

$$P(\mathbf{q}^*) = P(0, \ldots, 0, q_i^*, \ldots, q_{i+p}^*, 0, \ldots, 0) = P(\underbrace{0, \ldots, 0}_{(i+p'-2) \text{ times}}, q_{i+p'-1}, q_{i+p'}, 0, \ldots, 0).$$

Finally if Case (iii) occurs, then $g_1(\nu) > g_2(\nu) > \cdots > g_k(\nu)$. By Corollary 4.8(ii), $q_j^* = 0$ for $j < k$. So $P(\mathbf{q}^*) = P(0, \ldots, 0, 1)$.

The discussions in the above paragraph describe the possible forms of $P(\mathbf{q}^*)$. As a direct consequence, we have

(4.18) $$M^{\vec{\gamma}} = \min\{R_1, \ldots, R_{k-1}, P(0, \ldots, 0, 1)\},$$

where

(4.19) $\quad R_i := \min\{P(q_1, \ldots, q_k) : (q_1, \ldots, q_k) \in \Delta_k, \ q_j = 0 \text{ for } j \in \{1, \ldots, k\} \setminus \{i, i+1\}\}$

for $i = 1, \ldots, k - 1$. In what follows, we further investigate the values of R_i.

Fix $i \in \{1, \ldots, k - 1\}$. By definition,

(4.20) $$R_i = \inf_{t \in [0,1]} P(\underbrace{0, \ldots, 0}_{(i-1) \text{ times}}, 1 - t, t, 0, \ldots, 0).$$

For each $\nu_{i+1} \in \mathcal{B}_{i+1}$,

$$\sup\{H(\nu) + \nu(\widetilde{\phi}) : \nu \in \mathcal{B}_1, \ \Pi_{i+1}\nu = \nu_{i+1}\} = H(\nu_{i+1}) + \nu_{i+1}(\widetilde{\phi}_{i+1}),$$

where $\widetilde{\phi}_{i+1}$ is the potential on $\widetilde{X}_{i+1}$ defined by $\widetilde{\phi}_{i+1}(x) = \log \mathbb{E}(N_{x_1}^{(i+1)})$ (see Remark 4.4). This, combined with the definition of P, yields (noting that $\widetilde{\phi}_{i+1} = \phi_{i+1}/(\gamma_1 + \cdots + \gamma_{i+1})$, where ϕ_{i+1} was defined in (2.12))

$$P(\underbrace{0,\ldots,0}_{(i-1)\text{ times}}, 1-t, t, 0, \ldots, 0)$$

$$= \sup_{\nu \in \mathcal{B}_1} \left\{ (1-t) \left[(\gamma_1 + \ldots + \gamma_i)(H(\nu) + \nu(\widetilde{\phi})) + \sum_{j=i+1}^{k} \gamma_j H(\Pi_j \nu) \right] \right.$$

$$\left. + t \left[(\gamma_1 + \ldots + \gamma_{i+1})(H(\nu) + \nu(\widetilde{\phi})) + \sum_{j=i+2}^{k} \gamma_j H(\Pi_j \nu) \right] \right\}$$

$$= \sup_{\nu_{i+1} \in \mathcal{B}_{i+1}} \left[(\gamma_1 + \ldots + \gamma_i + t\gamma_{i+1})\nu_{i+1}(\widetilde{\phi}_{i+1}) + (\gamma_1 + \ldots + \gamma_{i+1})H(\nu_{i+1}) \right.$$

$$\left. + \sum_{j=i+2}^{k} \gamma_j H(\Pi_{i+1,j}\nu_{i+1}) \right]$$

$$= P_{i+1}\left(\frac{\gamma_1 + \ldots + \gamma_i + t\gamma_{i+1}}{\gamma_1 + \ldots + \gamma_{i+1}} \right),$$

where $\Pi_{i+1,j} = \pi_{j-1} \circ \cdots \circ \pi_{i+1}$, and P_{i+1} is defined as in (2.13). So by (4.20) we have

$$R_i = \min\{ P_{i+1}(t) : (\gamma_1 + \ldots + \gamma_i)/(\gamma_1 + \ldots + \gamma_{i+1}) \leq t \leq 1 \}, \quad i = 1, \ldots, k-1.$$

Notice that $P(0, \ldots, 0, 1) = P_k(1)$. Recalling that $\widetilde{\theta}_i = (\gamma_1 + \ldots + \gamma_{i-1})/(\gamma_1 + \ldots + \gamma_i)$ for $2 \leq i \leq k$, the previous discussion yields

(4.21)
$$M^{\vec{\gamma}} = \min \left(\min_{2 \leq i \leq k} \inf\{ P_i(\theta) : \theta \in [\widetilde{\theta}_i, 1]\}, P_k(1) \right).$$

Finally we prove that

$$M^{\vec{\gamma}} = \begin{cases} P_{i_0}(\theta_{i_0}) & \text{if } i_0 \leq k \\ P_k(1) & \text{if } i_0 = k+1 \end{cases},$$

where we recall that i_0 and θ_{i_0} where defined just before the statement of Theorem 2.3.

Below we divide our discussions into 3 steps :

Step 1. First assume that $i_0 = 2$, i.e., $P_2'(1) \geq 0$. By the convexity of P_2 we see that $P_2'(\theta)$ is non-decreasing, so there are only two possibilities: (1) $P_2'(\widetilde{\theta}_2) \geq 0$ (in which case, $\theta_{i_0} = \widetilde{\theta}_2$) ; or (2) $P_2'(\theta_{i_0}) = 0$.

Recall $P_2'(\theta) = (\gamma_1 + \gamma_2) \cdot \nu_{\theta\phi_2}(\phi_2)$, where $\nu_{\theta\phi_2} \in \mathcal{B}_2$ denotes the equilibrium state of the function P_2 at θ.

If (1) occurs, then by the above derivative formula, we have

$$\nu_{\widetilde{\theta}_2\phi_2}(\phi_2) = (\gamma_1 + \gamma_2)\nu_{\widetilde{\theta}_2\phi_2}(\widetilde{\phi}_2) \geq 0.$$

To simplify the notation, we write $\nu_2 := \nu_{\widetilde{\theta}_2\phi_2}$. Let ν be the unique element in $\mathcal{B}_1$ such that $\Pi_2 \nu = \nu_2$ and $H(\nu) + \nu(\widetilde{\phi}) = H(\nu_2) + \nu_2(\widetilde{\phi}_2)$. Since $\nu_2(\widetilde{\phi}_2) \geq 0$, we have

(4.22)
$$H(\nu) + \nu(\widetilde{\phi}) \geq H(\nu_2) \geq H(\Pi_3 \nu) \geq \ldots \geq H(\Pi_k \nu).$$

In the meantime,

$$P_2(\widetilde{\theta}_2) = \gamma_1 \nu_2(\widetilde{\phi}_2) + (\gamma_1 + \gamma_2)H(\nu_2) + \sum_{j=3}^{k} \gamma_j H(\Pi_{2,j}\nu_2)$$

$$= \gamma_1(H(\nu) + \nu(\widetilde{\phi})) + \sum_{j=2}^{k} \gamma_j H(\Pi_j \nu)) = g_1(\nu).$$

By (4.22), we have $g_1(\nu) \le g_2(\nu) \le \cdots \le g_k(\nu)$. So by (4.12),

$$P_2(\widetilde{\theta}_2) = g_1(\nu) = \min\{g_1(\nu), \ldots, g_k(\nu)\} \le M^{\vec{\gamma}}.$$

But $P_2(\widetilde{\theta}_2) \ge \inf_{\mathbf{q} \in \Delta_k} P(\mathbf{q}) = M^{\vec{\gamma}}$. So we get $P_2(\widetilde{\theta}_2) = M^{\vec{\gamma}}$.

Next assume that Case (2) occurs. We still let ν_2 denote the equilibrium state of P_2 at θ_{i_0}. The derivative formula gives that $\nu_2(\widetilde{\phi}_2) = 0$ and so

(4.23)
$$P_2(\theta_{i_0}) = (\gamma_1 + \gamma_2)H(\nu_2) + \sum_{j=3}^{k} \gamma_j H(\Pi_{2,j}\nu_2).$$

Let ν be the unique element in $\mathcal{B}_1$ such that $\Pi_2 \nu = \nu_2$ and $H(\nu) + \nu(\widetilde{\phi}) = H(\nu_2) + \nu_2(\widetilde{\phi}_2) = H(\nu_2)$. Then $g_1(\nu) = g_2(\nu) \le \cdots \le g_k(\nu)$. By (4.23) and the fact that $\nu_2(\widetilde{\phi}_2) = 0$,

$$P_2(\theta_{i_0}) = \gamma_1 \nu_2(\widetilde{\phi}_2) + (\gamma_1 + \gamma_2)H(\nu_2) + \sum_{j=3}^{k} \gamma_j H(\Pi_{2,j}\nu_2) = g_1(\nu).$$

Similar to the discussion for Case (1), we also obtain $P_2(\theta_{i_0}) = M^{\vec{\gamma}}$.

It follows from Remark 4.1, as well as Corollary 4.8(ii) that in any of the cases discussed above, μ_{ν_2} is the unique Mandelbrot measure μ supported on K such that $\dim_H(\mu) = M^{\vec{\gamma}}$, conditional on $K \ne \emptyset$.

Step 2. Assume that $2 < i_0 \le k$. Then $P'_{i_0-1}(1) < 0$ and $P'_{i_0}(1) \ge 0$. However, notice that $P_{i_0-1}(1) = P_{i_0}(\widetilde{\theta}_{i_0})$ (both of them are equal to

$$P(\underbrace{0, \ldots, 0}_{(i_0-2) \text{ times}}, 1, 0, 0, \ldots, 0).$$

Let ν_{i_0-1} and ν_{i_0} denote the equilibrium states of P_{i_0-1} at 1 and P_{i_0} at $\widetilde{\theta}_{i_0}$, respectively. Assume at the moment that $\pi_{i_0-1}\nu_{i_0-1} = \nu_{i_0}$; we will check it at the end of the proof. Now from $P'_{i_0-1}(1) < 0$ it follows that $\nu_{i_0-1}(\widetilde{\phi}_{i_0-1}) < 0$. Next we consider two possible cases: (a) $P'_{i_0}(\widetilde{\theta}_{i_0}) > 0$; (b) $P'_{i_0}(\widetilde{\theta}_{i_0}) \le 0$. If Case (a) occurs, then $\nu_{i_0}(\widetilde{\phi}_{i_0}) > 0$ and $\theta_{i_0} = \widetilde{\theta}_{i_0}$. Let ν be the unique element in $\mathcal{B}_1$ such that $\Pi_{i_0}\nu = \nu_{i_0}$ and $H(\nu) + \nu(\widetilde{\phi}) = H(\nu_{i_0}) + \nu_{i_0}(\widetilde{\phi}_{i_0})$. We will check that $H(\nu) + \nu(\widetilde{\phi}) = H(\nu_{i_0-1}) + \nu_{i_0-1}(\widetilde{\phi}_{i_0-1})$ as well. Hence $H(\nu_{i_0}) < H(\nu) + \nu(\widetilde{\phi}) < H(\nu_{i_0-1})$. It follows that

$$g_1(\nu) \ge \ldots \ge g_{i_0-1}(\nu) > g_{i_0}(\nu) \le \ldots \le g_k(\nu).$$

However $P_{i_0}(\widetilde{\theta}_{i_0}) = g_{i_0}(\nu)$ so $P_{i_0}(\widetilde{\theta}_{i_0}) = \min\{g_i(\nu) : 1 \le i \le k\} \le M^{\vec{\gamma}}$. Hence we have $P_{i_0}(\theta_{i_0}) = P_{i_0}(\widetilde{\theta}_{i_0}) = M^{\vec{\gamma}}$.

Suppose Case (b) holds. Then $P'_{i_0}(\theta_{i_0}) = 0$. Following the similar argument as in Case (2) of Step 1, we see that $P_{i_0}(\theta_{i_0}) = M^{\vec{\gamma}}$ by considering ν defined as in Case (a). Note

that $\nu_{\theta_{i_0}\phi_{i_0}}(\widetilde{\phi}_{i_0}) = 0$. We have $P_{i_0}(\theta_{i_0}) = g_{i_0}(\nu)$ and $g_1(\nu) \geq g_2(\nu) \geq \cdots \geq g_{i_0-1}(\nu) = g_{i_0}(\nu) \leq g_{i_0+1}(\nu) \leq \cdots \leq g_k(\nu)$. So $P_{i_0}(\theta_{i_0}) = M^{\vec{\gamma}}$.

Like in the first step, in any case, $\mu_{\nu_{i_0}}$ is the unique Mandelbrot measure μ supported on K such that $\dim_H(\mu) = M^{\vec{\gamma}}$, conditional on $K \neq \emptyset$.

Step 3. Assume that $i_0 = k+1$. Then $P_i'(1) < 0$ for every $i \in \{2, \ldots, k\}$. It follows that for each $i \in \{1, \ldots, k-1\}$, $P(\underbrace{0, \ldots, 0}_{(i-1) \text{ times}}, 1-t, t, 0, \ldots, 0)$ is strictly decreasing in t. Hence by (4.18)-(4.20), $M^{\vec{\gamma}} = P(0, \ldots, 0, 1) = P_k(1)$. Let ν_k be the equilibrium state of P_k at 1. and ν be the unique element of $\mathcal{B}_1$ such that $\Pi_k \nu = \nu_k$ and $H(\nu) + \nu(\widetilde{\phi}) = H(\nu_k) + \nu_k(\widetilde{\phi}_k)$. Then $P_k(1) = g_k(\nu)$. We conclude as in the two first steps.

The proof of Theorem 4.5 will be complete after we have checked that in Step 2 we have $\pi_{i_0-1}\nu_{i_0-1} = \nu_{i_0}$ and $H(\nu) + \nu(\widetilde{\phi}) = H(\nu_{i_0-1}) + \nu_{i_0-1}(\widetilde{\phi}_{i_0-1})$.

To see that these equalities do hold, we look at the optimisation problem consisting in determining the equilibrium state ν_{i_0-1} of P_{i_0-1} at 1, by conditioning on the knowledge of $\pi_{i_0-1}\nu_{i_0-1}$: using the relation $\phi_{i_0-1} = \widetilde{\gamma}_{i_0-1}\widetilde{\phi}_{i_0-1}$, we have

$$P_{i_0-1}(1)$$

$$= \sup_{\rho_{i_0-1} \in \mathcal{B}_{i_0-1}} \widetilde{\gamma}_{i_0-1}\big(\rho_{i_0-1}(\widetilde{\phi}_{i_0-1}) + H(\rho_{i_0-1})\big) + \sum_{i=i_0}^{k} \gamma_i H(\Pi_{i_0-1,i}\rho_{i_0-1})$$

$$= \sup_{\rho_{i_0} \in \mathcal{B}_{i_0}} \widetilde{\gamma}_{i_0} H(\rho_{i_0}) + \sum_{i=i_0+1}^{k} \gamma_i H(\Pi_{i_0,i}\rho_{i_0})$$

$$+ \sup_{(p(\cdot|\widehat{b}))_{\widehat{b} \in \widetilde{\mathcal{A}}_{i_0}}} \sum_{\widehat{b} \in \mathcal{A}_{i_0}} \rho_{i_0}([\widehat{b}]) \cdot \widetilde{\gamma}_{i_0-1} \sum_{\substack{b \in \widetilde{\mathcal{A}}_{i_0-1} \\ \pi_{i_0-1}(b)=\widehat{b}}} -p(b|\widehat{b}) \log p(b|\widehat{b}) + p(b|\widehat{b}) \log \mathbb{E}(N_b^{(i_0-1)}),$$

where $(p(b|\widehat{b}))_{b \in \mathcal{A}_{i_0-1}, \pi_{i_0-1}(b)=\widehat{b}}$ is the probability vector such that $p(b|\widehat{b}) = \rho_{i_0-1}(b)/\rho_{i_0}(\widehat{b})$ when $\rho_{i_0}(\widehat{b}) > 0$, and any probability vector otherwise. Optimising over the families $(p(b|\widehat{b}))_{b \in \mathcal{A}_{i_0-1}, \pi_{i_0-1}(b)=\widehat{b}}$ yields

$$p(b|\widehat{b}) = \frac{\mathbb{E}(N_b^{(i_0-1)})}{\mathbb{E}(N_{\widehat{b}}^{(i_0)})}$$

when $\rho_{i_0}(\widehat{b}) > 0$, and

$$P_{i_0-1}(1) = \sup_{\rho_{i_0} \in \mathcal{B}_{i_0}} \widetilde{\gamma}_{i_0-1}\rho_{i_0}(\widetilde{\phi}_{i_0}) + \widetilde{\gamma}_{i_0} H(\rho_{i_0}) + \sum_{i=i_0+1}^{k} \gamma_i H(\Pi_{i_0,i}\rho_{i_0})$$

$$= \sup_{\rho_{i_0} \in \mathcal{B}_{i_0}} \rho_{i_0}(\widetilde{\theta}_{i_0}\phi_{i_0}) + \widetilde{\gamma}_{i_0} H(\rho_{i_0}) + \sum_{i=i_0+1}^{k} \gamma_i H(\Pi_{i_0,i}\rho_{i_0})(= P_{i_0}(\widetilde{\theta}_{i_0})),$$

from which the desired equalities follow. $\qquad\square$

4.3. Lower bound for the Hausdorff dimension of K.
The sharp lower bound comes from the optimisation problem solved in Section 4.2. Consider the unique Mandelbrot

measure $\mu = \mu_{\nu_{\theta_{i_0} \phi_{i_0}}}$ obtained there. By construction the measure μ is fully supported on K conditional on $K \neq \emptyset$, because $\nu_{\theta_{i_0} \phi_{i_0}}$ is fully supported on $\widetilde{X}_{i_0}$ if $i_0 \leq k$ and $\widetilde{X}_k$ otherwise, and due to the last claim of Remark 4.2. Also, the assumptions of Theorem 2.2 are fulfilled for μ, and the Hausdorff dimension of μ provides the desired lower bound for $\dim_H K$.

4.4. **Upper bound for the Hausdorff dimension of K.** Let us start by discussing a first possible attempt to show that $\dim_H K \leq M^{\vec{\gamma}}$. We could expect to use the measure $\mu = \mu_{\nu_{\theta_{i_0} \phi_{i_0}}}$ of maximal Hausdorff dimension $M^{\vec{\gamma}}$ and show that $\underline{\dim}_{\mathrm{loc}}(\mu, x) \leq M^{\vec{\gamma}}$ everywhere on K; this is the approach used by McMullen [44] as well as Kenyon and Peres [38] in the deterministic case; that would make it possible to conclude quite quickly. In the random situation, we can show that this approach via the lower local dimension works in the case when $N_b^{(2)} \geq 2$ almost surely for all $b \in \widetilde{\mathcal{A}}_2$; say in this case that K is of type I. This requires quite involved moments estimates for martingales in varying environments. Notice that in this case we have $i_0 = 2$ and $\theta_2 = \gamma_1/(\gamma_1 + \gamma_2)$. The type I makes it possible to treat the case of a slightly more general type of examples, still quite close to the deterministic case: $i_0 = 2$, $\theta_2 = \gamma_1/(\gamma_1 + \gamma_2)$, and it is possible to approximate K by a sequence $(K^{(p)})_{p \in \mathbb{N}}$ of random Sierpinski sponges of type I in the sense that $K \subset K^{(p)}$ for all $p \in \mathbb{N}$, $\bigcap_{p \in \mathbb{N}} K^{(p)} = K$, and $\lim_{p \to \infty} \dim_H K^{(p)} = M^{\vec{\gamma}}$. A sufficient condition to be in this situation is that $\Psi_2(0) < \Psi_2(\gamma_1/(\gamma_1 + \gamma_2))$, where $\Psi_2(\theta) = \sum_{b \in \mathcal{A}_2} \mathbb{E}(N_b^{(2)})^\theta$ (this condition obviously holds for examples of type I).

Thus, regarding the lower local dimension approach, it remains open whether or not in general it holds that $\underline{\dim}_{\mathrm{loc}}(\mu, x) \leq M^{\vec{\gamma}}$ everywhere on K; moreover, the sufficient condition just stated to get the sharp upper bound for $\dim_H K$ is not at all satisfactory.

The alternative is to examine the strategy that Gatzouras and Lalley adopted for the two dimensional case. Their approach is inspired by Bedford's treatment of the deterministic two dimensional case, and it uses effective coverings of the set K to find the sharp upper bound for $\dim_H K$. These coverings are closely related to a combinatoric argument due to Bedford. But this argument turns out to be hard to extend to higher dimensional cases. Below, we use a different, though related, combinatoric argument, which yields nice effective coverings as well, but works in any dimension. Also, in the deterministic case and in any dimension, it provides an alternative to the argument using a uniform bound for the lower local dimension of μ. However, and interestingly, our argument uses a slight generalisation of a key combinatoric lemma established by Kenyon and Peres to get this uniform bound.

We now provide a general upper bound for $\dim_H K$, expressed through a variational principle.

Theorem 4.10. *With probability 1, conditional on $K \neq \emptyset$,*
$$\dim_H K \leq \inf \left\{ P_i(\theta) : \ i \in I, \ \widetilde{\theta}_i \leq \theta \leq 1 \right\}.$$

The fact that this is a sharp upper bound for $\dim_H K$ follows from (4.21) and the inequality $\dim_H K \geq M^{\vec{\gamma}}$ obtained in the previous section.

Before proving Theorem 4.10, we need to introduce some new definitions, and to make some preliminary observations.

Let $2 \leq i \leq k$ and $\widetilde{\theta}_i \leq \theta \leq 1$. For $\rho_i \in \mathcal{B}_i$ set

$$(4.24) \qquad D_{i,\theta}(\rho_i) = \widetilde{\gamma}_i \theta \sum_{b \in \mathcal{A}_i} \rho_i([b]) \log \mathbb{E}(N_b^{(i)}) + \widetilde{\gamma}_i h_{\rho_i}(T_i) + \sum_{j=i+1}^{k} \gamma_j h_{\Pi_{i,j}\rho_i}(T_j)$$

(again with the convention $0 \times (-\infty) = 0$.

For $\rho = (\widetilde{\rho}_i, \rho_i, \ldots, \rho_k) \in \mathcal{B}_i \times \prod_{j=i}^{k} \mathcal{B}_j$, set

$$(4.25) \qquad \widetilde{D}_{i,\theta}(\rho) = \widetilde{\gamma}_i \theta \sum_{b \in \mathcal{A}_i} \widetilde{\rho}_i([b]) \log \mathbb{E}(N_b^{(i)}) + \widetilde{\gamma}_i h_{\rho_i}(T_i) + \sum_{j=i+1}^{k} \gamma_j h_{\rho_j}(T_j).$$

For each $2 \leq i \leq k$. We endow the set $\mathcal{B}_i \times \prod_{j=i}^{k} \mathcal{B}_j$ with the distance

$$d_i(\rho, \rho') = \max \left(\max_{b \in \mathcal{A}_i} |\widetilde{\rho}_i([b]) - \widetilde{\rho}'_i([b])|, \ \max_{i \leq j \leq k} \max_{b \in \mathcal{A}_j} |\rho_j([b]) - \rho'_j([b])| \right),$$

which makes it a compact set. Let

$$(4.26) \qquad \mathcal{R}_i = \left\{ \rho \in \mathcal{B}_i \times \prod_{j=i}^{k} \mathcal{B}_j : \ \widetilde{D}_{i,\theta}(\rho) \leq P_i(\theta) \right\}.$$

The set $\mathcal{R}_i$ is compact. For any $\epsilon \in (0,1)$, $\mathcal{R}_i$ can be covered by a finite collection of open balls $\{\mathring{B}(\rho^{(m)}, \epsilon)\}_{1 \leq m \leq M(\epsilon)}$. Moreover, if $\epsilon \leq \min\{(\#\mathcal{A}_j)^{-1} : \ 1 \leq j \leq k\}$, we can assume that for all m the components of each probability vector $\rho^{(m)}$ are not smaller than $\epsilon/2$.

For $x \in X_1$, $2 \leq j \leq k$, and $n \in \mathbb{N}^*$ we define $\rho_j(x,n)$ to be the Bernoulli product measure on X_j associated with the probability vector whose components are the frequences of occurrence of the different elements of $\mathcal{A}_j$ in $\Pi_j(x_{|n})$, namely the vector

$$\left(n^{-1} \#\{1 \leq m \leq n : \ \Pi_j(x_m) = b\} \right)_{b \in \mathcal{A}_j}.$$

Also, let

$$(4.27) \qquad\qquad\qquad \rho(x,n) = (\widetilde{\rho}_i, \rho_i, \ldots, \rho_k),$$

where

$$\widetilde{\rho}_i = \rho_i \left(x, \lceil \theta \ell_i(n) \rceil \right), \ \rho_i = \rho_i(x, \ell_i(n)),$$
$$\rho_j = \rho_j \left(T_1^{\ell_{j-1}(n)} x, \ell_j(n) - \ell_{j-1}(n) \right), \quad j = i+1, \ldots, k.$$

Now, for any $n \in \mathbb{N}^*$ and $U = (U_i, \ldots, U_k) \in \mathcal{A}_i^{\ell_i(n)} \times \prod_{j=i+1}^{k} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$, we can define $\rho(U) = (\widetilde{\rho}_i(U), \rho_i(U), \ldots, \rho_k(U))$ as equal to $\rho(x,n)$, for any $x \in X_1$ such that $\Pi_i(x_1 \cdots x_{\ell_i(n)}) = U_i$ and $\Pi_j(x_{\ell_{j-1}(n)+1} \cdots x_{\ell_j(n)}) = U_j$ for all $i+1 \leq j \leq k$. Note that $\rho_i(U)$ depends on U_i only, so we also denote it by $\rho_i(U_i)$.

Then, for each $1 \leq m \leq M(\epsilon)$ and $n \in \mathbb{N}^*$ we set

$$\mathcal{R}_i(\epsilon, m, n) = \left\{ U \in \mathcal{A}_i^{\ell_i(n)} \times \prod_{j=i+1}^{k} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)} : \rho(U) \in \mathring{B}(\rho^{(m)}, \epsilon) \right\}.$$

It is standard to observe that if $U_i \in \mathcal{A}_i^{\ell_i(n)}$ is such that $|\rho_i(U_i)([b]) - \rho_i^{(m)}([b])| \leq \epsilon$ for all $b \in \mathcal{A}_i$ then

$$\rho_i^{(m)}([U_i]) = \prod_{b \in \mathcal{A}_i} \rho_i^{(m)}([b])^{\ell_i(n)\rho_i(U)([b])} \geq \prod_{b \in \mathcal{A}_i} \rho_i^{(m)}([b])^{\ell_i(n)\rho_i^{(m)}([b])} \prod_{b \in \mathcal{A}_i} \rho_i^{(m)}([b])^{\ell_i(n)\epsilon}$$

$$\geq \exp\left(-\ell_i(n)\left(h_{\rho_i^{(m)}}(T_i) + \epsilon \log(2/\epsilon)\right)\right),$$

Consequently, the cardinality of the set $\mathcal{U}_{i,\epsilon,m,n}$ of such U_i is bounded from above by $\exp\left(\ell_i(n)\left(h_{\rho_i^{(m)}}(T_i) + \epsilon \log(1/\epsilon)\right)\right)$.

Similarly, for each $i + 1 \leq j \leq k$, the cardinality of the set $\mathcal{U}_{j,\epsilon,m,n}$ of those $U_j \in \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$ such that $|\rho_j(U_j)([b]) - \rho_j^{(m)}([b])| \leq \epsilon$ for all $b \in \mathcal{A}_j$ is bounded from above by $\exp\left((\ell_j(n) - \ell_{j-1}(n))(h_{\rho_j^{(m)}}(T_j) + \epsilon \log(2/\epsilon))\right)$. Since by definition of $\mathcal{R}_i(\epsilon, m, n)$ we have $\mathcal{R}_i(\epsilon, m, n) \subset \prod_{i=j}^{k} \mathcal{U}_{j,\epsilon,m,n}$, the previous observations yield

$$\#\mathcal{R}_i(\epsilon, m, n)$$

$$\leq \prod_{j=i}^{k}(\#\mathcal{U}_{j,\epsilon,m,n})$$

(4.28) $\quad \leq \exp\left(\ell_k(n)\epsilon \log(2/\epsilon)\right) \exp\left(\ell_i(n)h_{\rho_i^{(m)}}(T_i) + \sum_{j=i+1}^{k}(\ell_j(n) - \ell_{j-1}(n))h_{\rho_j^{(m)}}(T_j)\right).$

We also notice that if we endow X_i with the metric

$$d_{\vec{\gamma},i}(x, y) = \max\left(e^{-\frac{|\Pi_{i,j}(x) \wedge \Pi_{i,j}(y)|}{\gamma_1 + \cdots + \gamma_j}} : j = i, \ldots, k\right),$$

the balls of radius $e^{-\frac{n}{\gamma_1}}$ in X_1 project to the balls of the same radius in X_i, which are parametrized by the elements of $\mathcal{A}_i^{\ell_i(n)} \times \prod_{j=i+1}^{k} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$, in the sense that such a ball takes the form

$$B_U = \left\{y \in X_i : y_1 \cdots y_{\ell_i(n)} = U_i, \ \Pi_{i,j}(y_{\ell_{j-1}(n)+1} \cdots y_{\ell_j(n)}) = U_j \text{ for } j = i+1, \ldots, k\right\}$$

for some U in $\mathcal{A}_i^{\ell_i(n)} \times \prod_{j=i+1}^{k} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$. Moreover, given such a ball B_U, $\Pi_i^{-1}(B_U) \cap K$ is covered by, say, a family $\mathcal{B}(U)$ of n_U balls of radius $e^{-\frac{n}{\gamma_1}}$ which intersect K. Each of the $N_{U_{i|\ell_{i-1}(n)}}^{(i)}$ cylinders of generation $\ell_{i-1}(n)$ in X_1 which intersects K and project to $[U_{i|\ell_{i-1}(n)}]$ in X_i via Π_i intersects only one such ball. Indeed, for such a cylinder $[V_1 \cdots V_{\ell_{i-1}}(n)]$, the data $\Pi_j([V_{\ell_{j-1}(n)+1} \cdots V_{\ell_j(n)}])$, $1 \leq j \leq i - 1$, and B_U determine a unique ball B of X_1 such that $\Pi_i(B) = B_U$. This implies $n_U \leq N_{U_{i|\ell_{i-1}(n)}}^{(i)}$. Consequently, for every integer ℓ between $\ell_{i-1}(n)$ and $\ell_i(n)$, we also have $n_U \leq N_{U_{i|\ell}}^{(i)}$. In particular,

(4.29) $$n_U \leq N_{U_{i|\lceil \theta \ell_i(n) \rceil}}^{(i)}.$$

The following lemma, whose proof we postpone to the end of this section, will play an essential role to find effective coverings of $\Pi_i(\widetilde{X}_1)$, and then of K. Let us mention at the moment that in this lemma $(1) \Rightarrow (2) \Rightarrow (3)$. ALso, recall the definition (4.26) of $\mathcal{R}_i$.

Lemma 4.11. *For all $x \in \widetilde{X}_1$:*

(1) $\liminf_{n\to\infty} \widetilde{D}_{i,\theta}(\rho(x,n)) - D_{i,\theta}(\rho_i(x,n)) \leq 0$;

(2) $\liminf_{n\to\infty} \widetilde{D}_{i,\theta}(\rho(x,n)) \leq P_i(\theta)$;

(3) *there exists $\rho \in \mathcal{R}_i$ and an increasing sequence of integers $(n_j)_{j\in\mathbb{N}}$ such that $\rho(x,n_j)$ converges to ρ as $j \to \infty$.*

Proof of Theorem 4.10. It follows from Lemma 4.11(3) that given $\epsilon > 0$, for all $x \in \widetilde{X}_1$, there exists $1 \leq m \leq M(\epsilon)$ such that $\Pi_i(x)$ belongs to $\bigcup_{U\in\mathcal{R}_i(\epsilon,m,n)} B_U$ for infinitely many integers n. As a result, for all $N \in \mathbb{N}^*$, we get the following covering of K:

$$K \subset \bigcup_{n\geq N} \bigcup_{m=1}^{M(\epsilon)} \bigcup_{U\in\mathcal{R}_i(\epsilon,m,n)} \bigcup_{B\in\mathcal{B}(U)} B.$$

Thus, given $s > 0$, the pre-Hausdorff measure $\mathcal{H}^s_{e^{-\frac{n}{\gamma_1}}}$ of K is bounded as follows:

$$\mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K) \leq \sum_{n\geq N} \sum_{m=1}^{M(\epsilon)} \sum_{U\in\mathcal{R}_i(\epsilon,m,n)} N^{(i)}_{U_{i|\lceil\theta\ell_i(n)\rceil}} e^{-\frac{n}{\gamma_1}s}.$$

Consequently, denoting by $(U_i)_\ell$ the ℓ-th letter of U_i,

$$\mathbb{E}\left(\mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K)\right) \leq \sum_{n\geq N} e^{-\frac{n}{\gamma_1}s} \sum_{m=1}^{M(\epsilon)} \sum_{U\in\mathcal{R}_i(\epsilon,m,n)} \mathbb{E}(N^{(i)}_{U_{i|\lceil\theta\ell_i(n)\rceil}})$$

$$= \sum_{n\geq N} e^{-\frac{n}{\gamma_1}s} \sum_{m=1}^{M(\epsilon)} \sum_{U\in\mathcal{R}_i(\epsilon,m,n)} \prod_{\ell=1}^{\lceil\theta\ell_i(n)\rceil} \mathbb{E}(N^{(i)}_{(U_i)_\ell})$$

and using the definition of $\widetilde{\rho}_i(U)$ to re-express the right hand side of the last inequality, we obtain

$$\mathbb{E}\left(\mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K)\right) \leq \sum_{n\geq N} e^{-\frac{n}{\gamma_1}s} \sum_{m=1}^{M(\epsilon)} \sum_{U\in\mathcal{R}_i(\epsilon,m,n)} \exp\left(\lceil\theta\ell_i(n)\rceil \sum_{b\in\mathcal{A}_i} \widetilde{\rho}_i(U)([b]) \log \mathbb{E}(N^{(i)}_b)\right).$$

Now we use the fact that $U \in \mathcal{R}_i(\epsilon,m,n)$ means that $d(\rho(U),\rho^{(m)}) \leq \epsilon$, to get a constant C_i independent of m, U and n such that

$$\sum_{b\in\mathcal{A}_i} \widetilde{\rho}_i(U)([b]) \log \mathbb{E}(N^{(i)}_b) \leq C_i\epsilon + \sum_{b\in\mathcal{A}_i} \widetilde{\rho}_i^{(m)}([b]) \log \mathbb{E}(N^{(i)}_b).$$

We then obtain:

$$\mathbb{E}\left(\mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K)\right) \leq \sum_{n\geq N} e^{-\frac{n}{\gamma_1}s} e^{\lceil\theta\ell_i(n)\rceil C_i\epsilon}$$

$$\cdot \sum_{m=1}^{M(\epsilon)} (\#\mathcal{R}_i(\epsilon,m,n)) \exp\left(\lceil\theta\ell_i(n)\rceil \sum_{b\in\mathcal{A}_i} \widetilde{\rho}_i^{(m)}([b]) \log \mathbb{E}(N^{(i)}_b)\right).$$

37

Using (4.28), the fact that $|\ell_j(n) - \frac{\gamma_1 + \ldots + \gamma_j}{\gamma_1} n| \leq 1$ for all $1 \leq j \leq k$, as well as the definition of $\widetilde{D}_{i,\theta}(\rho^{(m)})$, we deduce that there exists a constant $\widehat{C}_i$ such that for all $1 \leq m \leq M(\epsilon)$:

$$(\#\mathcal{R}_i(\epsilon, m, n)) \exp \left(\lceil \theta \ell_i(n) \rceil \sum_{b \in \mathcal{A}_i} \widetilde{\rho}_i^{(m)}([b]) \log \mathbb{E}(N_b^{(i)}) \right)$$

$$\leq \widehat{C}_i \exp \left(\ell_k(n) \epsilon \log(2/\epsilon) \right) \exp \left(\frac{n}{\gamma_1} \widetilde{D}_{i,\theta}(\rho^{(m)}) \right)$$

$$\leq \widehat{C}_i \exp \left(\ell_k(n) \epsilon \log(2/\epsilon) \right) \exp \left(\frac{n}{\gamma_1} P_i(\theta) \right) \quad \text{(recall that } \rho^{(m)} \in \mathcal{R}_i).$$

Upon taking $C_i = \widehat{C}_i$ big enough, we conclude that

$$\mathbb{E}\left(\mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K) \right) \leq C_i M(\epsilon) \sum_{n \geq N} \exp \left(-\frac{n}{\gamma_1} \left(s - P_i(\theta) - C_i \epsilon \log(2/\epsilon) \right) \right).$$

If $s > P_i(\theta) + C_i \epsilon \log(1/\epsilon)$, this yields $\mathbb{E}\left(\sum_{N \geq 1} \mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K) \right) < \infty$, so $\lim_{N \to \infty} \mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K) = 0$ and $\dim_H K \leq s$ almost surely. Since this holds for any given small enough $\epsilon > 0$, we get $\dim_H K \leq P_i(\theta)$ almost surely.

The previous upper bound is easily seen to hold simultaneously for all $2 \leq i \leq k$ and $\widetilde{\theta}_i \leq \theta \leq 1$ since its holds simultaneously for all $2 \leq i \leq k$ and rational $\widetilde{\theta}_i \leq \theta \leq 1$, and the mappings $\theta \mapsto P_i(\theta)$ are continuous. This yields Theorem 4.10. $\qquad \square$

Proof of Lemma 4.11. That $(1) \Rightarrow (2)$ follows from the fact that $P_i(\theta) = \max\{D_{i,\theta}(\nu_i) : \nu_i \in \mathcal{B}_i\}$, and $(2) \Rightarrow (3)$ is immediate.

Let $x \in X_1$. To prove (1), we are going to show that there exists $J \in \mathbb{N}^*$, as well as J bounded sequences $u_j : \mathbb{N}^* \to \mathbb{R}$ such that $\lim_{n \to \infty} u_j(n+1) - u_j(n) = 0$, and J couples $(\alpha_j, \beta_j) \in \mathbb{R}_+^*$ such that for all $n \in \mathbb{N}^*$,

$$(4.30) \qquad \widetilde{D}_{i,\theta}(\rho(x, n)) - D_{i,\theta}(\rho_i(x, n)) \leq \epsilon_n + \sum_{j=1}^J u_j(\lceil \beta_j n \rceil) - u_j(\lceil \alpha_j n \rceil),$$

with $\lim_{n \to \infty} \varepsilon_n = 0$. The desired conclusion is then a direct application of [26, Lemma 5.4], which is a slight extension of the combinatorial lemma used by Kenyon and Peres [38, Lemma 4.1].

To prove (4.30), noting that $\Pi_{i,j} \rho_i(x, n) = \rho_j(x, n)$ for all $i \leq j \leq k$, and using the respective definitions of $\widetilde{D}_{i,\theta}$ and $D_{i,\theta}$, we can write, after defining the sequences $v_i(n) = \widetilde{\gamma}_i \theta \sum_{b \in \mathcal{A}_i} \rho(x, n)([b]) \log \mathbb{E}(N_b^{(i)})$ and $w_j(n) = \gamma_j h_{\rho_j(x,n)}(T_j)$:

$$\widetilde{D}_{i,\theta}(\rho(x, n)) - D_{i,\theta}(\rho_i(x, n)) = v_i(\lceil \theta \ell_i(n) \rceil) - v_i(n) + \sum_{j=i}^k w_j(\ell_j(n)) - w_j(n)$$

$$+ \sum_{j=i+1}^k \gamma_j (h_{\rho_j(T_1^{\ell_{j-1}(n)} x, \ell_j(n) - \ell_{j-1}(n))}(T_j) - h_{\rho_j(x, \ell_j(n))}(T_j)).$$

38

Note that each $u \in \{v_i, w_i, \dots, w_k\}$ is bounded and does satisfy $\lim_{n\to\infty} u(n+1) - u(n) = 0$. Also, we have

$$h_{\rho_j(T_1^{\ell_{j-1}(n)}x, \ell_j(n) - \ell_{j-1}(n))}(T_j) - h_{\rho_j(x, \ell_j(n))}(T_j)$$

$$\text{(4.31)} \qquad \leq \frac{\ell_{j-1}(n)}{\ell_j(n) - \ell_{j-1}(n)}(h_{\rho_j(x, \ell_j(n))}(T_j) - h_{\rho_j(x, \ell_{j-1}(n))}(T_j)).$$

To see this, note that denoting $\rho_j(x, \ell_j(n))$, $\rho_j(x, \ell_{j-1}(n))$ and $\rho_j(T_1^{\ell_{j-1}(n)}x, \ell_j(n) - \ell_{j-1}(n))$ by ν_j, ν_j' and ν_j'' respectively, (4.31) is equivalent to

$$\text{(4.32)} \qquad (\ell_j(n) - \ell_{j-1}(n))\, h_{\nu_j''}(T_j) + \ell_{j-1}(n)\, h_{\nu_j'}(T_j) \leq \ell_j(n)\, h_{\nu_j}(T_j).$$

However, by definition, for all $b \in \mathcal{A}_j$ we have

$$\nu_j([b]) = \frac{\ell_{j-1}(n)}{\ell_j(n)}\nu_j'([b]) + \frac{\ell_j(n) - \ell_{j-1}(n)}{\ell_j(n)}\nu_j''([b]).$$

Thus, (4.32), and then (4.31), follow from the concavity of $x \geq 0 \mapsto -x\log(x)$ and the fact that $h_\nu(T_j) = -\sum_{b \in \mathcal{A}_j} \nu([b])\log([b])$ for $\nu \in \{\nu_j, \nu_j', \nu_j''\}$.

Setting $\alpha_j = \frac{\widetilde{\gamma}_j}{\gamma_1}$, (4.31) implies that

$$\gamma_j(h_{\rho_j(T_1^{\ell_{j-1}(n)}x, \ell_j(n) - \ell_{j-1}(n))}(T_j) - h_{\rho_j(x, \ell_j(n))}(T_j))$$

$$\leq \frac{\widetilde{\gamma}_{j-1}}{\gamma_j}(w_j(\lceil \alpha_j n \rceil) - w_j(\lceil \alpha_{j-1} n \rceil)) + o(1).$$

Moreover, $v_i(\lceil \theta \ell_i(n) \rceil) - v_i(n) = v_i(\lceil \theta \alpha_i n \rceil) - v_i(n) + o(1)$. Finally (4.30) holds. $\qquad\square$

5. The box counting dimension of K. Proofs of Theorem 2.4 and Corollary 2.6

Recall that for all $u \in \mathbb{N}^*$ and $u \in \mathcal{A}_1^*$, (K^u, K_n^u) denotes the copy of (K, K_n) generated by the random sets $(A(u \cdot a))_{a \in \mathcal{A}_1^*}$ (see (2.3)).

Proof of Theorem 2.4. Here again, without loss of generality we assume that all the γ_i are positive.

We will use in an essential way the result established in [27, Section 4], which deals with the case where $k = 2$, $m_1 = e^{-\gamma_1}$ and $m_2 = e^{-(\gamma_1 + \gamma_2)}$ are integers, and with the Euclidean realisation of K. It is worth noting that this result is strongly based on a result by Dekking on the asymptotic behaviour of the survival probability of a branching process in a random environment [14].

We first need to describe the balls of radius $e^{-\frac{n}{\gamma_1}}$ which intersect K. For $n \in \mathbb{N}^*$, we saw that the set $\mathcal{F}_n$ of balls in X_1 of radius $e^{-\frac{n}{\gamma_1}}$ equals the set $\{B_U : U = (U_1, \dots, U_k) \in \prod_{i=1}^k \mathcal{A}_i^{\ell_i(n) - \ell_{i-1}(n)}\}$, where

$$B_U = \{y \in X_1 : \Pi_i(T_1^{\ell_{i-1}(n)}(y))_{|\ell_i(n) - \ell_{i-1}(n)} = U_i, \ \forall 1 \leq i \leq k\}.$$

Thus $B_U \cap K \neq \emptyset$ if and only if the event

$$E_U = \left\{ \exists\, (u_i)_{1 \leq i \leq k} \in \prod_{i=1}^k \mathcal{A}_1^{\ell_i(n) - \ell_{i-1}(n)} : \ \text{both } F_k^U(u_1, \dots, u_k) \text{ and } K^{u_1 u_2 \cdots u_k} \neq \emptyset \text{ hold} \right\}$$

39

holds, where for all $1 \leq i \leq k$

$$F_i^{(U_1,U_2,\ldots,U_i)}(u_1,\ldots,u_i) = \left\{ \begin{array}{c} \Pi_j(u_j) = U_j, \ \forall\, 1 \leq j \leq i, \\ [u_j] \cap K_{\ell_j(n)-\ell_{j-1}(n)}^{u_1 u_2 \cdots u_{j-1}} \neq \emptyset, \ \forall\, 1 \leq j \leq i \end{array} \right\}$$

(note that necessarily $u_1 = U_1$).

For $2 \leq i \leq k$ and $(U_1,\ldots,U_i) \in \prod_{j=1}^{i} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$, we set

$$E_i(U_1,\ldots,U_i) = \left\{ \exists\, (u_j)_{1 \leq j \leq i} \in \prod_{j=1}^{i} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)} : \ F_i^{(U_1,U_2,\ldots,U_i)}(u_1,\cdots,u_i) \text{ holds} \right\}$$

(note that $E_1(U_1)$ is simply the event $\{[U_1] \cap K_n \neq \emptyset\}$). We deduce from [27, Section 4] that conditional on $K \neq \emptyset$, we have

$$\lim_{n\to\infty} \frac{\log \#\{(U_1,U_2) \in \mathcal{A}_1^n \times \mathcal{A}_2^{\ell_2(n)-n} : \ E_2(U_1,U_2) \text{ holds}\}}{n} = \log(\mathbb{E}(\#A)) + \frac{\gamma_2}{\gamma_1}\psi_2(\widehat{\theta}_2).$$

This result mainly comes from the fact that $\lim_{n\to\infty} \frac{\log \#\{U_1 \in \mathcal{A}_1^n : E_1(U_1) \text{ holds}\}}{n} = \log(\mathbb{E}(\#A)) > 0$, and given $U_1 \in \mathcal{A}_1^n$ such that $[U_1] \cap K_n \neq \emptyset$, the number of those $U_2 \in \mathcal{A}_2^{\ell_2(n)-n}$ such that $E_2(U_1,U_2)$ holds is a random variable $Z_{2,\ell_2(n)-n}(U_1)$, so that the random variables $Z_{2,\ell_2(n)-n}(U_1)$ are independent and identically distributed, $\lim_{n\to\infty} \frac{\log \mathbb{E}(Z_{2,\ell_2(n)-n})(U_1)}{n} = \frac{\gamma_2}{\gamma_1}\psi_2(\widehat{\theta}_2) > 0$ and, conditional on $K^{U_1} \neq \emptyset$, $\lim_{n\to\infty} \frac{\log Z_{2,\ell_2(n)-n}(U_1)}{n} = \frac{\gamma_2}{\gamma_1}\psi_2(\widehat{\theta}_2)$ almost surely.

Now for $2 \leq i \leq k$ set

$$s_i = \log(\mathbb{E}(\#A)) + \sum_{j=2}^{i} \frac{\gamma_i}{\gamma_1}\psi_j(\widehat{\theta}_j).$$

Suppose that $k \geq 3$, and for some $3 \leq i \leq k$ we have proven that conditional on $K \neq \emptyset$, it holds that

(5.1) $$\lim_{n\to\infty} \frac{\log \#\{(U_1,\ldots,U_{i-1}) \in \prod_{j=1}^{i-1} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)} : \ E_{i-1}(U_1,\ldots,U_{i-1}) \text{ holds}\}}{n} = s_{i-1}.$$

Given $(U_1,\ldots,U_{i-1}) \in \prod_{j=1}^{i-1} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$, and fixed associated $(u_1,\ldots,u_{i-1})$ such that $F_{i-1}^{(U_1,U_2,\ldots,U_{i-1})}(u_1,\cdots,u_{i-1})$ holds, following the arguments of [27], the cardinality of the set of those words $U_i \in \mathcal{A}_i^{\ell_i(n)-\ell_{i-1}(n)}$ such that there exists $u_i \in \mathcal{A}_1^{\ell_i(n)-\ell_{i-1}(n)}$ such that $F_i^{(U_1,U_2,\ldots,U_i)}(u_1,\cdots,u_i)$ holds, is a random variable $Z_{i,\ell_i(n)-\ell_{i-1}(n)}(U_1,\ldots,U_{i-1})$ so that the $Z_{i,\ell_i(n)-\ell_{i-1}(n)}(U_1,\ldots,U_{i-1})$ are independent, and identically distributed. Moreover, setting $\widetilde{\ell}_i(n) = \ell_i(n) - \ell_{i-1}(n)$, one has both $\lim_{n\to\infty} \frac{\log \mathbb{E}(Z_{i,\widetilde{\ell}_i(n)})(U_1,\ldots,U_{i-1})}{n} = \frac{\gamma_i}{\gamma_1}\psi_i(\widehat{\theta}_i) > 0$ and, conditional on $K^{u_1 \cdots u_{i-1}} \neq \emptyset$, $\lim_{n\to\infty} \frac{\log Z_{i,\widetilde{\ell}_i(n)}(U_1,\ldots,U_{i-1})}{n} = \frac{\gamma_i}{\gamma_1}\psi_i(\widehat{\theta}_i)$ almost surely. Then, again the same reasoning as in [27] for the case $k = 2$ with the roles of $\mathcal{A}_1^n$ and $\mathcal{A}_2^{\ell_2(n)-n}$ now respectively played by $\prod_{j=1}^{i-1} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$ and $\mathcal{A}_i^{\ell_i(n)-\ell_{i-1}(n)}$ shows that (5.1) holds for i as well. Consequently, applying this to $i = k$, conditional on $K \neq \emptyset$,

we get for all $n \geq 1$ an integer $\mathcal{N}_n$ such that $\lim_{n \to \infty} \frac{\log \mathcal{N}_n}{n} = s_k$, as well as $\mathcal{N}_n$ elements $U = (U_1, \ldots, U_k) \in \prod_{i=1}^k \mathcal{A}_i^{\ell_i(n) - \ell_{i-1}(n)}$ and associated $(u_1 = U_1, u_2, \ldots, u_k) \in \prod_{i=1}^k \mathcal{A}_1^{\ell_i(n) - \ell_{i-1}(n)}$ such that $F_k^{(U_1, \ldots, U_k)}(u_1, \ldots, u_k)$ hold. The events $\{K^{u_1 u_2 \cdots u_k} \neq \emptyset\}$ being independent, with the same probability $\mathbb{P}(K \neq \emptyset)$, and independent of the events $F_k^{(U_1, U_2, \ldots, U_k)}(u_1, u_2, \ldots, u_k)$, using [27, Section 4] again yields

$$\lim_{n \to \infty} \frac{\log \#\{U \in \prod_{i=1}^k \mathcal{A}_i^{\ell_i(n) - \ell_{i-1}(n)} : E_U \text{ holds}\}}{n} = \log(\mathbb{E}(\#A)) + \sum_{i=2}^k \frac{\gamma_i}{\gamma_1} \psi_i(\widehat{\theta}_i),$$

which, after dividing by γ_1^{-1}, is precisely $\lim_{n \to \infty} \frac{\log \#\{B \in \mathcal{F}_n : B \cap K \neq \emptyset\}}{-\log(e^{-\frac{n}{\gamma_1}})}$, i.e. $\dim_B K$. $\qquad \square$

Next we state, using our notation, a fact established in the proof of [7, Corollary 3.5], which is a variational approach to the dimension of projections of fractal percolation sets in a symbolic space $X_1 \times X_1$ to one of its two natural factors.

Proposition 5.1. Let $2 \leq i \leq k$. With probability 1, conditional on $K \neq \emptyset$,

$$\max \Big\{ \min(\dim_e(\mu), h_{\nu_i}(T_i)) : \mu \text{ is a Mandelbrot measure supported on } K \Big\} = \psi_i(\widehat{\theta}_i),$$

where ν_i stands for the expectation of $\Pi_i(\mu)$. Moreover the maximum is uniquely reached if and only if $\widehat{\theta}_i > 0$ or $\widehat{\theta}_i = 0$ and $\psi_i'(\widehat{\theta}_i) = 0$. In any case, when the maximum is reached, one has $\nu_i = \nu_{i, \widehat{\theta}_i}$, where

$$\nu_{i,\theta}([b]) = \mathbb{E}(N_b^{(i)})^\theta / \sum_{b' \in \mathcal{A}_i} \mathbb{E}(N_{b'}^{(i)})^\theta.$$

Also, if $\widehat{\theta}_i > 0$, or $\widehat{\theta}_i = 0$ and $\psi_i'(\widehat{\theta}_i) = 0$, then $\dim_e(\mu) = h_{\nu_i}(T_i)$ for the unique μ at which the maximum is reached, and if $\widehat{\theta}_i = 0$ and $\psi_i'(\widehat{\theta}_i) > 0$, then $\dim_e(\mu) > h_{\nu_i}(T_i)$ for μ at which the maximum is reached.

Proof of Corollary 2.6. It follows from (2.5) and Proposition 5.1 that for $\dim_H K = \dim_B(K)$ to hold almost surely, conditional on $K \neq \emptyset$, the Mandelbrot measure μ of maximal dimension supported on K must satisfy $\dim_e(\mu) = \log \mathbb{E}(\#A)$ and $\min(\dim_e(\mu), h_{\nu_i}(T_i)) = \psi_i(\widehat{\theta}_i)$ for all $i \in I$.

The condition $\dim_e(\mu) = \log \mathbb{E}(\#A)$ implies that μ is the so called branching measure, i.e. it is obtained from the random vector $(\mathbf{1}_A(a)/\mathbb{E}(\#A))_{a \in \mathcal{A}_1}$. The other condition implies that for all $2 \leq i \leq k$, we have $\nu_i = \nu_{i, \widehat{\theta}_i}$. Since μ is the branching measure, this implies that $\mathbb{E}(N_b^{(i)})^{\widehat{\theta}_i} / e^{\psi_i(\widehat{\theta}_i)} = \mathbb{E}(N_b^{(i)}) / \mathbb{E}(\#A)$ for all $b \in \widetilde{\mathcal{A}}_i$, hence $\mathbb{E}(N_b^{(i)})^{\widehat{\theta}_i - 1}$ does not depend on $b \in \widetilde{\mathcal{A}}_i$. This is a non trivial condition only if $\widehat{\theta}_i < 1$. This proves the necessity of the condition given in the statement.

Now assume that $\mathbb{E}(N_b^{(i)})$ does not depend on $b \in \widetilde{\mathcal{A}}_i$ for all $i \in I$ such that $\widehat{\theta}_i < 1$. Suppose first that there is no $i \in I$ such that $\widehat{\theta}_i < 1$, i.e. $\widehat{\theta}_i = 1$ for all $i \in I$. By the remark made above, the branching measure μ does satisfy $\dim_e^{\vec{\gamma}}(\mu) = (\gamma_1 + \cdots + \gamma_k) \dim_e(\mu) = \dim_B(K)$. Next, suppose that $\widehat{\theta}_i < 1$ for some $i \in I$. Again, consider the branching measure μ. Since $\mathbb{E}(N_b^{(i)})$ does not depend on $b \in \widetilde{\mathcal{A}}_i$ we do have $\nu_i = \nu_{i, \widehat{\theta}_i}$, so that $\min(\dim_e(\mu), h_{\nu_i}(T_i)) = h_{\nu_i}(T_i) = h_{\nu_{i, \widehat{\theta}_i}}(T_i) = \psi_i(\widehat{\theta}_i)$. This yields again $\dim_e^{\vec{\gamma}}(\mu) = \dim_B K$. $\qquad \square$

6. Projections of K and μ to factors of X_1. Proofs of Theorems 2.7, 2.8 and 2.10, and Corollary 2.11

The proofs will be sketched.

Sketch of the proof of Theorem 2.7. For all $n \geq 1$, denote by $\mathcal{F}_n^i$ the set of balls of X_i of radius $e^{-\frac{n}{\gamma_1}}$. Let $j_0 = \max\{i \leq j \leq k : T'(1) \leq \mathcal{T}'_{\nu_j}(1)\}$, with the convention $\max(\emptyset) = i - 1$. Computations similar to those used to prove Theorem 2.2 yield $q_0 > 1$ and $c_0 \geq 0$ such that for all $q \in (0, q_0]$ we have

$$\mathbb{E}\Big(\sum_{B \in \mathcal{F}_n^i} \Pi_i \mu(B)^q \Big) = O\big(\exp(-t(j_0, q, n)) \big) \quad \text{as } n \to \infty,$$

where

$$t(j_0, q, n) = \begin{cases} \ell_{j_0}(n)(T(q) - c_0(q-1)^2) + \sum_{j=j_0+1}^k (\ell_j(n) - \ell_{j-1}(n))\mathcal{T}_{\nu_j}(q) & \text{if } j_0 \geq i \\ \ell_i(n)\mathcal{T}_{\nu_i}(q) + \sum_{j=i+1}^k (\ell_j(n) - \ell_{j-1}(n))\mathcal{T}_{\nu_j}(q) & \text{otherwise.} \end{cases}$$

This is enough to get the differentiability of $\tau_{\Pi_i \mu}$ at 1 with $\tau'_{\Pi_i \mu}(1)$ equal to $\dim_e^{\vec{\gamma}^i}(\Pi_i \mu)$, and conclude. $\qquad\square$

Sketch of the proof of Theorem 2.8. We start with the lower bound.

If we proceed as in Section 4.5, we have to consider

$$M^{\vec{\gamma}^i} = \max \left\{ \sum_{j=i}^k \gamma_j^i \min\{H(\nu) + \nu(\widetilde{\phi}), H(\Pi_i \nu)\} : \nu \in \mathcal{B}_1 \right\}.$$

Define $\widetilde{g}_i : \mathcal{B}_1 \to \mathbb{R}$ ($\mathcal{B}_1$ being still identified with $\mathcal{P}(\widetilde{\mathcal{A}}_1)$) by

$$\widetilde{g}_i(\nu) = (\gamma_1 + \ldots + \gamma_i)H(\Pi_i \nu) + \sum_{j=i+1}^k \gamma_j H(\Pi_j \nu) = h_{\Pi_i \nu}^{\vec{\gamma}^i}(T_i).$$

Note that $\widetilde{g}_i$ is strictly concave only when considered as a function of $\Pi_i \nu$.

We have, remembering the definition (4.11) of the mappings g_j,

$$M^{\vec{\gamma}^i} = \sup_{\nu \in \mathcal{B}_1} \min(\widetilde{g}_i(\nu), g_i(\nu), \ldots, g_k(\nu)).$$

To express this supremum in terms of the pressure functions $P_j(\cdot)$, $i \leq j \leq k$, we have to consider the mapping $\widetilde{P} : \Delta_{k-i+2} \to \mathbb{R}$ defined by

$$\widetilde{P}(\widetilde{q}_i, q_i, \ldots, q_k) = \sup_{\nu \in \mathcal{B}_1} \widetilde{q}_i \widetilde{g}_i(\nu) + q_i g_i(\nu) + \ldots + q_k g_k(\nu).$$

Corollary 4.7 yields

$$M^{\vec{\gamma}^i} = \min_{(\widetilde{q}_i, q_i, \ldots, q_k) \in \Delta_{k-i+2}} \widetilde{P}(\widetilde{q}_i, q_i, \ldots, q_k).$$

Let $q^* = (\widetilde{q}_i^*, q_i^*, \ldots, q_k^*)$ at which the minimum is attained, and $\nu \in \mathcal{B}_1$ such that $\widetilde{P}(q^*) = \widetilde{q}_i^* \widetilde{g}_i(\nu) + q_i^* g_i(\nu) + \ldots + q_k^* g_k(\nu)$. Due to Remark 4.4, one has

$$\widetilde{P}(q^*) = \sup_{\nu_i \mathcal{B}_i} \widetilde{q}_i^* \widetilde{f}_i(\nu_i) + q_i^* f_i(\nu_i) + \ldots + q_k^* f_k(\nu_i),$$

where $\widetilde{f}_i(\nu_i) = h_{\nu_i}^{\vec{\gamma}^i}(T_i)$ and

$$f_p(\nu_i) = (\gamma_1 + \cdots + \gamma_p)(H(\nu_i) + \nu_i(\widetilde{\phi}_i)) + \sum_{j=p+1}^{k} \gamma_j H(\Pi_{i,j}\nu_i), \quad p = i, i+1, \ldots, k.$$

Moreover, the supremum is reached at a unique element ν_i of $\mathcal{B}_i$, which is fully supported. As in the minimization of $P(q_1, \ldots, q_k)$ in the proof of Theorem 4.5, either one of the following 3 cases occurs: (i) $H(\nu_i) + \nu_i(\widetilde{\phi}_i) \geq H(\nu_i)$; (ii) there exists a unique $j \in \{i, \ldots, k-1\}$ so that $H(\Pi_{i,j+1}\nu_i) \leq H(\nu_i) + \nu_i(\widetilde{\phi}_i) < H(\Pi_{i,j}\nu_i)$; (iii) $H(\nu_i) + \nu_i(\widetilde{\phi}_i) < H(\Pi_{i,k}\nu_i)$. In cases (ii) and (iii), we necessarily have $\widetilde{q}_i^* = 0$, and we conclude by using the same approach as in the study of P that either there exists $j' \in \{j, \ldots, k-1\}$ and $(q_{j'}, q_{j'+1}) \in \Delta_2$ such that $\widetilde{P}(\widetilde{q}_i^*, q_i^*, \ldots, q_k^*) = \widetilde{P}(\underbrace{0, \ldots, 0}_{(j'-i+1) \text{ times}}, q_{j'}, q_{j'+1}, 0, \ldots, 0)$, or

$q^* = (0, \ldots, 0, 1)$. Also, if $j' > i$, then

$$\widetilde{P}(\underbrace{0, \ldots, 0}_{(j'-i+1) \text{ times}}, q_{j'}, q_{j'+1}, 0, \ldots, 0) = P(\underbrace{0, \ldots, 0}_{(j'-i+1) \text{ times}}, q_{j'}, q_{j'+1}, 0, \ldots, 0).$$

Now, suppose that we are in case (i). Let j be the smallest index j' in $\{i, \ldots, k\}$ such that $f_{j'}(\nu_i) > \widetilde{P}(\widetilde{q}_i^*, q_i^*, \ldots, q_k^*)$, and let $j = k+1$ otherwise. We have $q_{j'}^* = 0$ for all $j \leq j' \leq k$. If $j \leq i+1$, we get $\widetilde{P}(\widetilde{q}_i^*, q_i^*, \ldots, q_k^*) = \widetilde{P}(\widetilde{q}_i^*, q_i^*, 0, \ldots, 0)$. If $j > i+1$, then $\pi_{j'}$ is injective for all $i \leq j' \leq j-2$. Set $\widetilde{\gamma}_s = \gamma_1 + \cdots + \gamma_s$. For each $(\widetilde{q}_i, q_i, \ldots, q_{j-1}) \in \Delta_{j-i+1}$, we have

$$\widetilde{P}(\widetilde{q}_i, q_i, \ldots, q_{j-1}, 0, \ldots, 0) = \sup_{\mu_i \in \mathcal{B}_i} \left(\sum_{t=i}^{j-1} q_t \widetilde{\gamma}_t \right) \mu_i(\phi_i) + \widetilde{\gamma}_i H(\mu_i) + \sum_{s=j}^{k} \gamma_s H(\Pi_{i,s}\mu_i),$$

where we have used that $H(\mu_i) = H(\Pi_{i,j'}\mu_i)$ for all $i < j' \leq j-1$ since $\pi_{j'}$ is injective for all $i \leq j' \leq j-2$. Thus $\widetilde{P}(\widetilde{q}_i, q_i, \ldots, q_{j-1}, 0, \ldots, 0)$ depends only on $\sum_{t=i}^{j-1} q_t \widetilde{\gamma}_t$. Set $\beta := \sum_{t=i}^{j-1} q_t^* \widetilde{\gamma}_t \in [0, \widetilde{\gamma}_{j-1}]$. Either $\beta \in [0, \widetilde{\gamma}_i]$ or $\beta \in [\widetilde{\gamma}_{j'}, \widetilde{\gamma}_{j'+1}]$ for some $i \leq j' \leq j-2$. If the first case occurs, then there exists $(\widetilde{q}_i, q_i) \in \Delta_2$ such that $q_i \widetilde{\gamma}_i = \beta$, so $\widetilde{P}(\widetilde{q}_i^*, q_i^*, \ldots, q_k^*) = \widetilde{P}(\widetilde{q}_i, q_i, 0, \ldots, 0)$. If the second case occurs, then there exists $(q_{j'}, q_{j'+1}) \in \Delta_2$ such that $\beta = q_{j'}\widetilde{\gamma}_{j'} + q_{j'+1}\widetilde{\gamma}_{j'+1}$, so

$$\widetilde{P}(\widetilde{q}_i^*, q_i^*, \ldots, q_k^*) = \widetilde{P}(\underbrace{0, \ldots, 0}_{(j'-i+1) \text{ times}}, q_{j'}, q_{j'+1}, 0, \ldots, 0).$$

Then, noting that $\inf_{t \in [0,1]} \widetilde{P}(t, 1-t, 0, \ldots, 0) = \inf_{t \in [0,1]} P_i(t)$, we get

$$M^{\vec{\gamma}^i} = \min\left(\inf\{P_i(\theta) : \theta \in [0,1]\}, \min_{i+1 \leq j \leq k} \inf\{P_j(\theta) : \theta \in [\widetilde{\theta}_i, 1]\}, P_k(1) \right).$$

Moreover, by adapting the discussion of the proof of Corollary 4.8, we can get that either $M^{\vec{\gamma}^i}$ equals $P_i(\theta)$ with $\theta \in [0,1]$ and in this case $M^{\vec{\gamma}^i}$ is necessarily attained by $\nu \in \mathcal{B}_1$ such that $\Pi_i \nu$ is the equilibrium state of P_i at θ, or there is a unique $\nu \in \mathcal{B}_1$ at which $M^{\vec{\gamma}^i}$ is attained; in the later case, the unique maximizing measure ν is the same as the maximizing measure for $M^{\vec{\gamma}}$ in Theorem 4.5 in case $i_0 \geq i+1$.

If $P_i'(1) \geq 0$, then, using the notations of the statement of Theorem 2.8, we have $j_0 = i$, and we denote by ν_i the equilibrium state of P_i at θ_i^i. Suppose that $P_i'(\theta_i^i) = 0$, which holds automatically if $\theta_i^i > 0$. This means that $\nu_i(\widetilde{\phi}_i) = 0$. Let ν be the unique

element of $\mathcal{B}_1$ such that $\Pi_i \nu = \nu_i$ and $H(\nu) + \nu(\widetilde{\phi}) = H(\nu_i) + \nu_i(\widetilde{\phi}_i) = H(\nu_i)$. We have $P_i(\theta_i^i) = \widetilde{g}_i(\nu) = g_i(\nu) \leq g_{i+1}(\nu) \leq \cdots \leq g_k(\nu)$, which implies $M^{\vec{\gamma}^i} = P_i(\theta_i^i)$, and due to the observation made in the previous paragraph and Remark 4.1, the measure $\mu_\nu = \mu_{\nu_{i_0}}$ is the unique Mandelbrot measure μ such that the dimension of $\Pi_i \mu$ equals $M^{\vec{\gamma}^i}$, conditional on $K \neq \emptyset$; moreover, one has $\dim_H(\mu_\nu) = \dim_H(\Pi_i \mu_\nu)$. If $\theta_i^i = 0$ and $P_i'(\theta_i^i) > 0$, we still have $M^{\vec{\gamma}^i} = P_i(\theta_i^i) = \widetilde{g}_i(\nu) = g_i(\nu) \leq g_{i+1}(\nu) \leq \cdots \leq g_k(\nu)$, but for any Mandelbrot measure μ supported on K, if $\rho = \mathbb{E}(\mu)$ satisfies $H(\rho) + \rho(\widetilde{\phi}) > H(\nu_i)$, then $\dim_H(\Pi_i \mu) = M^{\vec{\gamma}^i}$. There are infinitely many such measures. Indeed, μ_{ν_i} itself satisfies this property since $P_i'(0) = \nu_i(\widetilde{\phi}_i) = 0 > 0$. Then one gets other examples of measures by considering, as defined in Section 4.1, random vectors W written as the "skewed" product of $(\nu_i([b]))_{b \in \mathcal{A}_i}$ with random vectors $(V_b)_{b \in \mathcal{A}_i}$ obtained as slight perturbions of the vector $(\widetilde{V}_b)_{b \in \mathcal{A}_i}$ used to construct the measure μ_{ν_i}. If $P_i'(1) < 0$, the discussion ends like in the proof of Theorem 4.5.

For the upper bound for the Hausdorff dimension, we prove that

$$\dim_H \Pi_i(K) \leq \inf\{P_j(\theta) : \theta \in [0,1] \text{ if } j = i, \text{ and } \theta \in [\widetilde{\theta}_j, 1] \text{ if } i < j \leq k\},$$

which in view of the lower bound is enough to conclude. To show the previous inequality, we extend the definitions of $D_{i,\theta}$ and $\widetilde{D}_{i,\theta}$ (see (4.24) and (4.25)) to $\theta \in [0,1]$ and for $j = i$ we redefine the vector $\rho(x,n)$ of (4.27) by taking $\widetilde{\rho}_i = \rho_i(x, \ell_i(n))$. It is readily seen from the proof of Lemma 4.11 that the conclusions of this lemma is still valid with these new definitions of $\widetilde{D}_{i,\theta}$ and $\rho(x,n)$.

Now, arguing similarly as in the proof of Theorem 4.10, for each $j \in \{i, \ldots, k\}$, for each $U = (U_j, \cdots, U_k)$ in $\mathcal{A}_j^{\ell_j(n)} \times \prod_{j'=j+1}^{k} \mathcal{A}_{j'}^{\ell_{j'}(n) - \ell_{j'-1}(n)}$, $\Pi_{i,j}^{-1}(B_U) \cap \Pi_i(K)$ is covered by, say, a family $\mathcal{B}(U)$ of n_U balls of radius $e^{-\frac{n}{\gamma_1}} \subset X_i$ which intersect $\Pi_i(K)$.

Suppose $j = i$ and fix $\theta \in [0,1]$. In this case $n_U = 1$ and we can bound this number by $(N_{U_i}^{(i)})^\theta$. Noting that $\mathbb{E}((N_{U_i}^{(i)})^\theta) \leq \mathbb{E}(N_{U_i}^{(i)})^\theta$, we can use similar estimates as in the proof of Theorem 4.10 to now estimate $\mathcal{H}^s_{-\frac{N}{\gamma_1}}(\Pi_i(K)))$, and this yields $\dim_H \Pi_i(K) \leq P_i(\theta)$ (here we followed the same idea as that used in [19] to deal with projections of planar statistically self-similar limit sets of fractal percolation).

Next, suppose $j \in \{i+1, \ldots, k\}$ and fix $\theta \in [\widetilde{\theta}_j, 1]$. Denote by $C_{U_j}^{(i,j)}$ the set of cylinders of generation $\ell_{j-1}(n)$ in X_i which intersects $\Pi_i(K)$, and project to $[U_{j|\ell_{j-1}(n)}]$ in X_j via $\Pi_{i,j}$. Also denote by $N_{U_j}^{(i,j)}$ the cardinality of this set. Each cylinder in $C_{U_j}^{(i,j)}$ intersects at most one of the elements of B_U. Thus $n_U \leq N_{U_j}^{(i,j)}$, so that:

$$n_U \leq \sum_{b \in C_{U_j}^{(i,j)}} 1 \leq \sum_{b \in C_{U_j}^{(i,j)}} N_b^{(i)} = N_{U_{j-1}}^{(j)}.$$

Then, the same lines as in the proof of Theorem 4.10 yield $\dim_H \Pi_i(K) \leq P_i(\theta)$ for all $\theta \in [\widetilde{\theta}_j, 1]$. $\qquad\square$

Sketch of the proof of Theorem 2.10. This is similar to the proof of Theorem 2.4, except that one must evaluate the cardinality of those $B \in \mathcal{F}_n^i$ such that $B \cap \Pi_i(K) \neq \emptyset$, and

this time we exploit results known for the box dimension of projections of statistically self-similar fractal Euclidean percolation sets from dimension 2 to dimension 1.

We have to estimate the cardinality of those $U \in \mathcal{A}_i^{\ell_i(n)} \times \prod_{j=i+1}^{k} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$ such that $E^i(U)$ holds, with

$$E^i(U) = \Big\{ \exists\, (u_j)_{i \leq j \leq k} \in \mathcal{A}_1^{\ell_i(n)} \times \prod_{j=i+1}^{k} \mathcal{A}_1^{\ell_j(n)-\ell_{j-1}(n)} :$$

$$\text{both } F_k^{i,(U_i,\ldots,U_k)}(u_i,\ldots,u_k) \text{ and } K^{u_i u_2 \cdots u_k} \neq \emptyset \text{ hold}\},$$

and where

$$F_j^{i,(U_i,\ldots,U_j)}(u_i,\ldots,u_j) = \left\{ \begin{array}{l} \Pi_{j'}(u_{j'}) = U_{j'}, \ \forall\, i \leq j' \leq j, \\ [u_i] \cap K_{\ell_i(n)} \neq \emptyset \\ [u_{j'}] \cap K_{\ell_{j'}(n)-\ell_{j'-1}(n)}^{u_i \cdots u_{j'-1}} \neq \emptyset, \ \forall\, i+1 \leq j' \leq j \end{array} \right\}.$$

One deduces easily from [15] (see alternatively [19] or [7]), which deal with the case $k = 2$, that

$$\lim_{n \to \infty} \frac{\log \#\{U_i \in \mathcal{A}_i^{\ell_i(n)} : \exists u_i \in \mathcal{A}_1^{\ell_i(n)}, \ F_i^{i,U_i}(u_i) \text{ holds}\}}{n} = \frac{\widetilde{\gamma}_i}{\gamma_1} \psi_i(\widehat{\theta}_i).$$

Then, a recursion similar to that used in the proof of Theorem 2.4 yields the desired result

$$\lim_{n \to \infty} \frac{\log \#\{U \in \mathcal{A}_i^{\ell_i(n)} \times \prod_{j=i+1}^{k} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)} : \ E^i(U) \text{ holds}\}}{n} = \frac{\widetilde{\gamma}_i}{\gamma_1} \psi_i(\widehat{\theta}_i) + \sum_{j=i+1}^{k} \frac{\gamma_j}{\gamma_1} \psi_j(\widehat{\theta}_j),$$

i.e. $\dim_B \Pi_i(K) = \widetilde{\gamma}_i \psi_i(\widehat{\theta}_i) + \sum_{j=i+1}^{k} \gamma_j \psi_j(\widehat{\theta}_j)$ after normalizing by γ_1^{-1}.   □

Proof of Corollary 2.11. If $\widehat{\theta}_i = 1$, using Proposition 5.1 we see that the equality between $\dim_H \Pi_i(K)$ and $\dim_B \Pi_i(K)$ imposes that $\dim_H \Pi_i(K)$ is attained by the branching measure, and the situation boils down to that of Corollary 2.6. This gives point (1) of the statement.

Suppose now that $\dim_H \Pi_i(K) = \dim_B \Pi_i(K)$, $\widehat{\theta}_i < 1$ and $\psi_i'(\widehat{\theta}_i) = 0$ (which is automatically true if $0 < \widehat{\theta}_i < 1$). The equality between $\dim_H \Pi_i(K)$ and $\dim_B \Pi_i(K)$ imposes that if μ stands for the unique Mandelbrot measure supported on K such that $\dim_e^{\widetilde{\gamma}_i}(\mu) = \dim_H \Pi_i(K)$, then $\dim_e(\mu) = h_{\nu_i}(T_i) = \psi_i(\widehat{\theta}_i)$, where $\nu_i = \mathbb{E}(\Pi_i \mu) = \nu_{i,\widehat{\theta}_i}$. Also, for $j \in I_i (= \{i,\ldots,k\})$ such that $j \leq j_0'$, we must have $\Pi_{i,j} \nu_i = \nu_{j,\widehat{\theta}_j}$. Using that for all $b \in \widetilde{\mathcal{A}}_j$, we have $\sum_{b' \in \Pi_{i,j}^{-1}(b)} \nu_{i,\widehat{\theta}_i}([b']) = \nu_j([b])$ and the fact that $\psi_i'(\widehat{\theta}_i) = \psi_j'(\widehat{\theta}_j)$ we can write

$$0 = \psi_i'(\widehat{\theta}_i) - \psi_j'(\widehat{\theta}_j) = \sum_{b' \in \widetilde{\mathcal{A}}_i} \nu_{i,\widehat{\theta}_i}([b']) \log \mathbb{E}(N_{b'}^{(i)}) - \sum_{b \in \widetilde{\mathcal{A}}_j} \nu_{j,\widehat{\theta}_j}([b]) \log \mathbb{E}(N_b^{(j)})$$

$$= \sum_{b \in \widetilde{\mathcal{A}}_j} \sum_{b' \in \Pi_{i,j}^{-1}(b)} \nu_{i,\widehat{\theta}_i}([b']) \log \frac{\mathbb{E}(N_{b'}^{(i)})}{\mathbb{E}(N_b^{(j)})}.$$

This implies that for all $b \in \widetilde{\mathcal{A}}_j$, the set $\Pi_{i,j}^{-1}(b) \cap \widetilde{\mathcal{A}}_i$ is a singleton $\{b'\}$ such that $\mathbb{E}(N_{b'}^{(i)}) = \mathbb{E}(N_b^{(j)})$, hence $\psi_i = \psi_j$ and $\widehat{\theta}_j = \widehat{\theta}_i$. Let us now examine those $j > j_0'$ in I_i. The previous argument shows that $\widehat{\theta}_j = 0$ and $\psi_j'(0) > 0$ (for otherwise j_0' would be at least

45

equal to the smallest of those j), and $\Pi_{i,j}\nu_i = \nu_{j,0}$, so that $\Pi_{i,j}\nu_i$ is uniformly distributed, i.e. $\sum_{b' \in \Pi_{i,j}^{-1}(b)} \mathbb{E}(N_{b'}^{(i)})^{\widehat{\theta}_i}$ does not depend on $b \in \widetilde{\mathcal{A}}_j$. We conclude that the conditions of point (2) are necessary. Conversely, if these conditions hold, one easily checks using Theorems 2.8 and 2.10 that $\dim_H \Pi_i(K) = \dim_B \Pi_i(K)$.

The last case easily follows from the previous discussion. $\qquad\square$

7. Dimension of conditional measures. Proof of Theorem 2.13

Since the proof of Theorem 2.14 is very similar to that of Theorem 2.13, we leave it to the reader.

Point (3) of the statement simply follows from points (1) and (2) as well as the dimensions formula provided by Theorems 2.2 and 2.7 for $\dim(\mu)$ and $\dim(\Pi_i\mu)$.

To get point (1) we notice that for any $x \in X_1$ and $n \geq 1$ we have $[x_{|\ell_k(n)}] \subset B(x, e^{-n/\gamma_1}) \subset [x_{|n}]$, so since $\dim_e(\mu) \leq h_{\nu_i}(T_i)$ we find that Theorem 2.12 implies that μ^z is exact dimensional with Hausdorff dimension equal to 0.

Now we prove point (2). The following lines do not depend on $\Pi_i\mu$ being absolutely continuous with respect to ν_i or not.

When $\mu_\omega = \mu \neq 0$, for $\Pi_i\mu_\omega$-almost every z, the conditional measure μ_ω^z is supported on $K^z = \pi^{-1}(\{z\}) \cap K$, obtained as the weak-star limit, as $n \to \infty$, of the measures $\mu_{\omega,n}^z$ obtained on K by assigning uniformly the mass $\frac{\mu_\omega([J] \cap \Pi_i^{-1}([z_{|n}]))}{\Pi_i\mu_\omega([z_{|n}])}$ to each cylinder $[J]$ of generation n in X_1. To be more specific, for any cylinder $[J]$, almost surely, the measurable set

$$A_J = \left\{ (\omega, z) \in \Omega \times X_i : \lim_{n \to \infty} \frac{\mu_\omega([J] \cap \Pi_i^{-1}([z_{|n}]))}{\Pi_i\mu_\omega([z_{|n}])} \text{ exists} \right\}$$

is of full $\widehat{\mathbb{Q}}$-probability, where we define $\widehat{\mathbb{Q}}(\mathrm{d}\omega, \mathrm{d}z) = \mathbb{P}(\mathrm{d}\omega)\Pi_i\mu_\omega(\mathrm{d}z)$, and for all (ω, z) in a subset A'_J of A_J of full $\widehat{\mathbb{Q}}$-probability, we have $\mu_\omega^z([J]) = \lim_{n \to \infty} \frac{\mu_\omega([J] \cap \Pi_i^{-1}([z_{|n}]))}{\Pi_i\mu_\omega([z_{|n}])}$.

Suppose now that conditional on $\mu \neq 0$, $\Pi_i\mu_\omega$ is absolutely continuous with respect to ν_i. There exists a measurable set A' of full $\widehat{\mathbb{Q}}$-probability such that for all $(\omega, z) \in A'$, the limit

$$\lim_{n \to \infty} \left(f_{\omega,n}(z) := \frac{\Pi_i\mu_\omega([z_{|n}])}{\nu_i([z_{|n}])} \right)$$

exists and is positive. We denote it by $f_\omega(z)$.

Set $A = A' \cap \bigcap_{J \in \Sigma} A'_J$. For all $(\omega, z) \in A$, the sequence of measures $\widetilde{\mu}_{\omega,n}^z = f_{\omega,n}(z)\mu_{\omega,n}^z$ weakly converges to the measure $\widetilde{\mu}_\omega^z$ defined as $f_\omega(z)\mu_\omega^z$.

Let

$$\Omega_A = \{\omega : (\omega, z) \in A \text{ for some } z \in X_i\},$$
$$F^\omega = \{z \in X_i : (\omega, x) \in A\}, \quad \forall \omega \in \Omega_A.$$

If $(\omega, x) \notin A$, set $\mu_\omega^x = \widetilde{\mu}_\omega^x = 0$.

For $z \in F^\omega$, $n \geq 1$ and $J \in \mathcal{A}_1^n$, we have

$$\widetilde{\mu}^z([J]) = \lim_{p \to \infty} \widetilde{\mu}_p^z([J]) = \lim_{p \to \infty} \frac{\mu([J] \cap \Pi_i^{-1}([z_{|n+p}]))}{\nu_i([z_{|n+p}])}.$$

If $U = (U_1, \cdots, U_k) \in \prod_{i=1}^k \mathcal{A}_i^{\ell_i(n) - \ell_{i-1}(n)}$, the ball B_U intersects $\Pi_i^{-1}(\{z\})$ if and only if $\Pi_{j,i}(U_j) = T_i^{\ell_{j-1}(n)}(z)_{|\ell_j(n) - \ell_{j-1}(n)}$ for all $1 \leq j \leq i-1$ and $U_j = \Pi_{i,j}(T_i^{\ell_{j-1}(n)}(z)_{|\ell_j(n) - \ell_{j-1}(n)})$ for $i \leq j \leq k$. Recalling (3.3), we also have

$$\widetilde{\mu}^z(B_U) = \sum_{(J_1,\ldots,J_k) \in \mathcal{J}_U} \widetilde{\mu}^z([J_1 \cdots J_k]])$$

$$= \sum_{(J_1,\ldots,J_k) \in \mathcal{J}_U} \lim_{p \to \infty} \frac{\mu([J_1 \cdots J_k] \cap \Pi_i^{-1}([z_{|n+p}]))}{\nu_i([z_{|n+p}])}.$$

Fix $q \geq 0$. For all $n \geq 1$, we are going to estimate the expectation of the partition function $\sum_{B \in \mathcal{F}_n} \widetilde{\mu}^z(B_U)^q$ with respect to the measure $\mathbb{P} \otimes \nu_i$.

Let $j_0 = \min\{2 \leq j \leq i-1 : \dim_e(\mu) > h_{\nu_j}(T_j)\}$, with $\min(\emptyset) = i$, and

$$D = \widetilde{\gamma}_{j_0-1}(\dim_e(\mu) - h_{\nu_i}(T_i)) + \sum_{j=j_0}^{i-1} \gamma_j(h_{\nu_j}(T_j) - h_{\nu_i}(T_i)),$$

which is precisely the value given by (2.22) due to our choice of j_0. We will show that there exists $c > 0$ such that for all q in a neighbourhood of 1, there exists $C_q > 0$ such that we have

$$\mathbb{E}_{\mathbb{P} \otimes \nu_i}\Big(\sum_{B_U \in \mathcal{F}_n} \widetilde{\mu}^z(B_U)^q \Big) \leq C_q \exp\Big(- \frac{n}{\gamma_1}(q-1)D + O((q-1)^2)) \Big).$$

This is enough to conclude that with probability 1, conditional on $\mu \neq 0$, for $\Pi_i\mu$-almost every z (remember that $\Pi_i\mu$ is absolutely continuous with respect to ν_i), one has $\tau_{\widetilde{\mu}^z}(q)) \geq (q-1)D - c(q-1)^2$ in some neighbourhood of 1. But since μ^z is a multiple of $\widetilde{\mu}^z$, the same holds for μ^z. This implies that the concave functions τ_{μ^z} and $q \mapsto (q-1)D - c(q-1)^2$ share the same derivative at 1, namely D. Consequently, μ^z is exact dimensional with dimension D.

Now we prove (7.1). Recall that outside the set A, the measure $\widetilde{\mu}_\omega^z$ has been defined equal to 0. By Fatou's lemma, we have

$$\mathbb{E}_{\mathbb{P} \otimes \nu_i} \sum_{B_U \in \mathcal{F}_n} \widetilde{\mu}^z(B_U)^q$$

$$\leq \liminf_{p \to \infty} \mathbb{E}_{\mathbb{P} \otimes \nu_i} \sum_{B_U \in \mathcal{F}_n} \Big(\sum_{(J_1,\ldots,J_k) \in \mathcal{J}_U} \frac{\mu([J_1 \cdots J_k] \cap \Pi_i^{-1}([z_{|\ell_k(n)+p}]))}{\nu_i([z_{|\ell_k(n)+p}])} \Big)^q$$

$$= \liminf_{p \to \infty} \mathbb{E} \sum_{L \in \widetilde{\mathcal{A}}_i^{\ell_k(n)+p}} \sum_{B_U \in \mathcal{F}_n} \Big(\sum_{(J_1,\ldots,J_k) \in \mathcal{J}_U} \frac{\mu([J_1 \cdots J_k] \cap \Pi_i^{-1}([L]))}{\nu_i([L])} \Big)^q \nu_i([L])$$

$$= \liminf_{p \to \infty} \mathbb{E} \sum_{L \in \widetilde{\mathcal{A}}_i^{\ell_k(n)+p}} \sum_{B_U \in \mathcal{F}_n} \Big(\sum_{(J_1,\ldots,J_k) \in \mathcal{J}_U} \sum_{J_p' \in \Pi_i^{-1}(T_i^{\ell_k(n)}L)} \frac{\mu([J_1 \cdots J_k J_p'])}{\nu_i([L])} \Big)^q \nu_i([L]).$$

47

Denote by S the expectation in the right hand side of the previous inequality. Due to the remark made above about the condition for B_U to intersects $\Pi_i^{-1}([u])$, and the multiplicativity property of the measure ν_i, we can rewrite S as follows:

$$S = \mathbb{E} \sum_{L=(L_1,\ldots,L_{i-1})} \sum_{(U_1,\ldots,U_{i-1})_L} \sum_{L'} m(U_1,\ldots,U_{i-1},L,L')^q \nu_i([L_1 \cdots L_{i-1} L']),$$

where $L \in \prod_{j=1}^{i-1} \widetilde{\mathcal{A}}_i^{\ell_{i-1}(n)}$, $(U_1,\ldots,U_{i-1})_L \in \prod_{j=1}^{i-1} \widetilde{\mathcal{A}}_j^{\ell_j(n)-\ell_{j-1}(n)}$ is such that $\Pi_{j,i}(U_j) = L_j$ for each $1 \le j \le i-1$, $L' \in \widetilde{\mathcal{A}}_i^{p+\ell_k(n)-\ell_{i-1}(n)}$, and taking the conventions that the words involved below whose writing uses the symbol J belong to $\widetilde{\mathcal{A}}_1^*$,

$$m(U_1,\ldots,U_{i-1},L,L') = \sum_{\substack{(J_1,\ldots,J_{i-1}): \Pi_{j,i}(J_j)=U_j, \\ J=J_i \cdots J_k: \Pi_i(J_j)=L_j}} \sum_{J': \Pi_i(J')=L'} \frac{\mu([J_1 \cdots J_{i-1} J'])}{\nu_i([L_1 \cdots L_{i-1} L'])}.$$

Suppose that $q \ge 1$. Using the same idea as in the proof of Theorem 3.2, but rewriting S as an expectation with respect to $\mathbb{P} \otimes \nu_i$ instead of $\mathbb{P}$, yields

$$\mathbb{E}(S) \le 2^{3q(i-1)} \left(\prod_{j=1}^{i-1} S_{j,n} \right) \cdot R_{n,p},$$

where

$$S_{j,n} = \mathbb{E}\Big(\sum_{L_j} \nu_i([L_j]) \sum_{U_j: \Pi_{j,i}(U_j)=L_j} \Big(\sum_{J_j: \Pi_j(J_j)=U_j} \frac{\mu_{\ell_j(n)-\ell_{j-1}(n)}([J_j])}{\nu_i(L_j)} \Big)^q \Big)$$

and

$$R_{n,p} = \mathbb{E}\Big(\sum_{L'} \Big(\sum_{J': \Pi_i(J')=L'} \frac{\mu([J'])}{\nu_i([L'])} \Big)^q \nu_i([L']) \Big).$$

Note that $R_{n,p} = \mathbb{E}_{\mathbb{P} \otimes \nu_i}(X_{p+\ell_k(n)-\ell_{i-1}(n)}^q)$, where

$$X_n(\omega,z) = \sum_{J \in \widetilde{\mathcal{A}}_1^n: \Pi_i(J)=z_{|n}} \frac{\mu_\omega([J])}{\nu_i([z_{|n}])}$$

is a perturbation of the martingale in random environment

$$\widetilde{X}_n(\omega,z) = \sum_{J \in \widetilde{\mathcal{A}}_i^n: \Pi_i(J)=z_{|n}} \frac{\mu_{\omega,n}([J])}{\nu_i([z_{|n}])}.$$

Now, recalling the definition of the vectors $V_b^{(i)}$ in Section 4.1 and setting

$$\varphi(q) = \log \sum_{b \in \widetilde{\mathcal{A}}_1} \nu_i([b]) e^{-T_{V_b^{(i)}}(q)},$$

our assumption that $\dim_e(\mu) > h_{\nu_i}(T_i)$ is equivalent to saying that at point 1 the function φ has a negative derivative, since $\varphi'(1) = h_{\nu_i}(T_i) - T'(1) = h_{\nu_i}(T_i) - \dim_e(\mu)$. We can then apply [7, Proposition 5.1] to $X_{p+\ell_k(n)-\ell_{i-1}(n)}$ and for q close enough to $1+$, get a constant $C_q > 0$ such that $R_{p,n} \le C_q$ independently of n and p.

48

Next we estimate the terms $S_{j,n}$, for $1 \le j \le i - 1$. For $j = 1$, we simply have

$$S_{1,n} = \mathbb{E}\Big(\sum_{L_1} \nu_i([L_1]) \sum_{U_1 \,:\, \Pi_i(U_1)=L_1} \Big(\frac{\mu_n([U_1])}{\nu_i([L_1])}\Big)^q \Big) = e^{n\varphi(q)} = e^{n(q-1)(h_{\nu_i}(T_i)-\dim_e(\mu))+O((q-1)^2))}.$$

For $2 \le j \le i - 1$, we rewrite $S_{j,n}$ as (recall that ν_j stands for the expectation of $\Pi_j \mu$)

$$S_{j,n} = \mathbb{E} \sum_{U_j} \phi_j(U_j) \Big(\sum_{J_j \,:\, \Pi_j(J_j)=U_j} \frac{\mu_{\ell_j(n)-\ell_{j-1}(n)}([J_j])}{\nu_j([U_j])} \Big)^q,$$

where

$$\phi_j(U_j) = \nu_j([U_j])^q \nu_i(\Pi_{j,i}([U_j]))^{1-q}.$$

Let $\nu_{q,j}$ be the Bernoulli product measure on $\widetilde{X}_j$ associated with the probability vector $\nu_{q,j}([b]) = \dfrac{\phi_j([b])}{\sum_{b' \in \widetilde{\mathcal{A}}_j} \phi_j([b'])}$, and define

$$\varphi_j(q) = \log \sum_{b \in \widetilde{\mathcal{A}}_j} \phi_j([b]) \quad and \quad \widetilde{X}_n^{(j)}(\omega, z) = \sum_{J \in \widetilde{\mathcal{A}}_1^n \,:\, \Pi_j(J)=z_{|n}} \frac{\mu_n([J])}{\nu_j([z_{|n}])}, \ z \in \widetilde{X}_j.$$

With these definitions, $S_{j,n}$ rewrites

$$S_{j,n} = e^{(\ell_j(n)-\ell_{j-1}(n))\varphi_j(q)} \mathbb{E}_{\mathbb{P}\otimes\nu_{q,j}}((\widetilde{X}_n^{(j)})^q),$$

Again, we can use [7, Proposition 5.1], and get a constant $C_{q,j} > 0$ such that

$$S_{j,n} \le C_{q,j} e^{(\ell_j(n)-\ell_{j-1}(n))\varphi_j(q)} \max\Big(1, \sum_{b \in \widetilde{\mathcal{A}}_j} \nu_{q,j}([b])e^{-T_{V_b^{(j)}}(q)}\Big)^{\ell_j(n)-\ell_{j-1}(n)}.$$

A computations shows that the derivative at 1 of the function $q \mapsto \sum_{b \in \widetilde{\mathcal{A}}_j} \nu_{q,j}([b])e^{-T_{V_b^{(j)}}(q)}$ is equal to $h_{\nu_j}(T_j) - T'(1) = h_{\nu_j}(T_j) - \dim_e(\mu)$.

Recall that $j_0 = \min\{2 \le j \le i - 1 : \dim_e(\mu) > h_{\nu_j}(T_j)\}$, with $\min(\emptyset) = i$. If $j_0 \le j \le i - 1$ and q is close enough to 1, we thus have $\sum_{b \in \widetilde{\mathcal{A}}_j} \nu_{q,j}([b])e^{-T_{V_b^{(j)}}(q)} < 1$, hence $S_{j,n} \le C_{q,j} e^{(\ell_j(n)-\ell_{j-1}(n))\varphi_j(q)}$. If $2 \le j < j_0$, using a Taylor expansion of order 2 we get $\sum_{b \in \widetilde{\mathcal{A}}_j} \nu_{q,j}([b])e^{-T_{V_b^{(j)}}(q)} \le \exp((q-1)(h_{\nu_j}(T_j) - \dim_e(\mu)) + O((q-1)^2))$. Moreover, for any j, $e^{\varphi_j(q)} = \exp((q-1)(h_{\nu_i}(T_i) - h_{\nu_j}(T_j)) + O((q-1)^2))$. So for $2 \le j < j_0$, $S_{j,n} \le C_{q,j} \exp((\ell_j(n) - \ell_{j-1}(n))((q-1)(h_{\nu_i}(T_i) - \dim_e(\mu)) + O((q-1)^2))$. Finally, for q close enough to 1+, there exists $C_q > 0$ such that

$$\mathbb{E}_{\mathbb{P}\otimes\nu_i} \sum_{B_U \in \mathcal{F}_n} \widetilde{\mu}^z(B_U)^q \le C_q \exp\Big((q-1)(\ell_{j_0-1}(n)(h_{\nu_i}(T_i) - \dim_e(\mu))$$

$$+ (q-1)\sum_{j=j_0}^{i-1}(\ell_j(n) - \ell_{j-1}(n))(h_{\nu_i}(T_i) - h_{\nu_j}(T_j)) + O((q-1)^2)n\Big)$$

(7.1)
$$= C_q \exp\Big(-\frac{n}{\gamma_1}(q-1)D + O((q-1)^2))\Big),$$

hence (7.1) holds.

Suppose now that $q \in (0,1)$. Using the same idea as in the proof of Theorem 3.2 yields

$$\mathbb{E}(S) \leq \prod_{j=1}^{i-1} \widetilde{S}_{j,n},$$

where

$$\widetilde{S}_{j,n} = \begin{cases} \mathbb{E}\Big(\sum_{L_j} \nu_i([L_j]) \sum_{U_j : \Pi_{j,i}(U_j)=L_j} \Big(\sum_{J_j : \Pi_j(J_j)=U_j} \dfrac{\mu_{\ell_j(n)-\ell_{j-1}(n)}([J_j])^q}{\nu_i(L_j)^q} \Big) \Big) & \text{if } 1 \leq j \leq j_0 - 1, \\[2ex] \mathbb{E}\Big(\sum_{L_j} \nu_i([L_j])^{1-q} \sum_{U_j : \Pi_{j,i}(U_j)=L_j} \nu_j(U_j)^q \Big) & \text{if } j_0 \leq j \leq i - 1. \end{cases}$$

With the notations introduced in the case $q \geq 1$, this rewrites

$$\widetilde{S}_{j,n} = \begin{cases} \Big(\sum_{b \in \widetilde{\mathcal{A}}_i} \nu_i([b]) e^{-T_{V_b^{(i)}(q)}} \Big)^{\ell_j(n)-\ell_{j-1}(n)} & \text{if } 1 \leq j \leq j_0 - 1, \\[2ex] e^{\varphi_j(q)(\ell_j(n)-\ell_{j-1}(n))} & \text{if } j_0 \leq j \leq i - 1. \end{cases}$$

Using Taylor expansions we can get that (7.1) holds for q close to $1-$ as well.

8. The case when $\{2 \leq i \leq k : \gamma_i \neq 0\} = \emptyset$

In our main statements about the Hausdorff and box-counting dimension of K and its projections for simplicity we assumed all the γ_i, $2 \leq i \leq k$ to be positive, which in the Euclidean realisation of Section 2.6 corresponds to $m_1 > \cdots > m_k \geq 2$. It turns out that up to slight modifications in the statement and proofs, our results cover the general configuration $m_1 \geq \cdots \geq m_k \geq 2$, for which the diagonal endomorphism $\mathrm{diag}(m_1, \ldots, m_k)$ may have eigenspaces of dimension at least 2 over which it is a similarity. In this case, in the expressions giving the dimensions of K and its projections, when $m_i = m_{i-1}$, i.e. $\gamma_i = \frac{1}{\log(m_{i-1})} - \frac{1}{\log(m_i)} = 0$, the index i has no contribution, and geometrically for any $1 \leq i < j \leq k$, $x \in X_i$ and $n \geq 1$, for the induced metric by $d_{\vec{\gamma}}$ on X_i, if $y \in B(x, e^{-n/\gamma_1})$, nothing is required on $T_i^{\ell_{j-1}(n)}(y)_{|\ell_j(n)-\ell_{j-1}(n)}$.

For all the statements of Section 2.3 and Theorem 4.10, the only change to make to cover the case $\gamma_i \geq 0$ for all $2 \leq i \leq k$ is to set $I = \{2 \leq i \leq k : \gamma_i > 0$ and replace k by $\sup(I)$ in (2.18). The proofs adapt readily.

For the statements of Section 2.4, one has to replace I_i by $\{i\} \cup \{i < j \leq k : \gamma_j > 0\}$ and replace k by $\sup(I_i)$ in (2.21). Again, the modifications in the proofs are left to the reader.

Appendix A. Main notation and conventions

For the reader's convenience, we summarize in Table 1 the main notation and typographical conventions used in this paper.

TABLE 1. Main notation and conventions

$\mathbb{N}, \mathbb{N}^*, \mathbb{R}_+, \mathbb{R}_+^*$	sets of non-negative (resp. positive) integers, non-negative (resp. positive) real numbers
(X_i, T_i)	one-sided full shift space over a finite alphabet $\mathcal{A}_i$, $i = 1, \ldots, k$
$\mathcal{M}(X_i, T_i)$	set of T_i-invariant Borel probability measures on X_i, $i = 1, \ldots, k$
$\pi_0 : X_1 \to X_1$	identity map on X_1
$\pi_i : X_i \to X_{i+1}$	one-block factor map, $i = 1, \ldots, k-1$
$\Pi_i : X_1 \to X_i$	$\Pi_i = \pi_{i-1} \circ \cdots \circ \pi_0$
$\Pi_{i,j} : X_i \to X_j$	$\Pi_{i,j} = \pi_{j-1} \circ \cdots \circ \pi_i$ if $j > i$, and the identity map if $j = i$
$\vec{\gamma} = (\gamma_1, \ldots, \gamma_k)$	a fixed vector in $\mathbb{R}_+^* \times (\mathbb{R}_+)^{k-1}$
$\vec{\gamma}^i = (\vec{\gamma}_j^i)_{i \le j \le k}$	the vector $(\gamma_1 + \ldots + \gamma_i, \gamma_{i+1}, \ldots, \gamma_k)$
$d_{\vec{\gamma}}$	an ultrametric distance defined on X_1 (cf. (2.1))
$(X_1, d_{\vec{\gamma}})$	self-affine symbolic space
$K = K(\omega)$	symbolic random statistically self-affine sponge defined from a random subset A of $\mathcal{A}_1$, where A is alternatively written as $(c_a)_{a \in \mathcal{A}_1}$ (Section 2.1)
$W = (W_a)_{a \in \mathcal{A}_1}$	a non-negative random vector such that $W_a > 0 \subset \{c_a = 1\}$ and $\mathbb{E}(\sum_{a \in \mathcal{A}_1} W_a) = 1$ (Section 2.2)
$(A(u), W(u))_{u \in \mathcal{A}_1^*}$	independent copies of (A, W) (Section 2.2)
μ	Mandelbrot measure associated with W (Section 2.2)
K_μ	topological support of a Mandelbrot measure μ
$\dim_e(\nu)$	entropy dimension of a finite Borel measure on X_i (cf. 2.4)
$h_\nu(T_i)$	measure-theoretic entropy of T_i with respect to a T_i-invariant measure ν
$\dim_e^{\vec{\gamma}}(\mu)$	$\sum_{i=1}^k \gamma_i \dim_e(\Pi_i \mu) = \gamma_1 \dim_e(\mu) + \sum_{i=2}^k \gamma_i \min(\dim_e(\mu), h_{\nu_i}(T_i))$ (cf. (2.6))
$h_\nu^{\vec{\beta}}(T_i)$	$\sum_{j=i}^k \beta_j h_{\Pi_{i,j}\nu}(T_j)$, where $\vec{\beta} = (\beta_i, \ldots, \beta_k)$ and $\nu \in \mathcal{M}(X_i, T_i)$
$N_b^{(i)}$	$\#\{a \in \mathcal{A}_1 : [a] \subset \Pi_i^{-1}([b]), [a] \cap K \ne \emptyset\}$, $b \in \mathcal{A}_i$
$\widetilde{\mathcal{A}_i}, \widetilde{X_i}$	$\widetilde{\mathcal{A}_i} = \{b \in \mathcal{A}_i : \mathbb{E}(N_b^{(i)}) > 0\}$, and $\widetilde{X_i}$ is the one-sided full shift space over $\widetilde{\mathcal{A}_i}$
$P^{\vec{\beta}}(\phi, T_i)$	weighted topological pressure of $\phi \in C(\widetilde{X_i})$ (cf. (2.10))
$P_i(\theta)$	$P^{\vec{\gamma}^i}(\theta\phi_i, T_i)$, where $\phi_i(x) = (\gamma_1 + \cdots + \gamma_i) \log \mathbb{E}(N_{x_1}^{(i)})$, $i = 2, \ldots, k$
i_0, θ_{i_0}	(cf. (2.16) and (2.17))
ψ_i	$\log \sum_{b \in \widetilde{\mathcal{A}_i}} \mathbb{E}(N_b^{(i)})^\theta$, $i = 2, \ldots, k$
$\widehat{\theta}_i$	(cf. the definition right after (2.19))

REFERENCES

[1] K. Barański. Hausdorff dimension of the limit sets of some planar geometric constructions. *Adv. Math.*, **210** (2007), 215–245.

[2] B. Bárány, A. Käenmäki, H. Koivusalo. Dimension of self-affine sets for fixed translation vectors. *J. Lond. Math. Soc.*, **98** (2018), 223–252.

[3] B. Bárány, M. Hochman, A. Rapaport. Hausdorff dimension of planar self-affine sets and measures. *Invent. Math.*, **216** (2019), 601–659.

[4] B. Bárány, M. Rams, K. Simon. On the dimension of triangular self-affine sets. *Ergod. Th. & Dynam. Sys.*, (2019), 1751–1783.

[5] J. Barral, D. J. Feng. Non-uniqueness of ergodic measures with full Hausdorff dimension on Gatzouras-Lalley carpet. *Nonlinearity*, **24** (2011), 2563–2567.

[6] J. Barral, D. J. Feng. Weighted thermodynamic formalism on subshifts and applications. *Asian J. Math.*, **16** (2012), 319–352.

[7] J. Barral, D. J. Feng. Projections of random Mandelbrot measures. *Adv. Math.*, 325 (2018), 640–718.

[8] T. Bedford. Crinkly curves, Markov partitions and box dimension in self-similar sets. Ph.D. Thesis, University of Warwick, 1984.

[9] F. Ben Nasr. Dimension de Hausdorff de certains fractals aléatoires. *Sém. Théor. Nombres Bordeaux*, **4** (1992), 129–140.

[10] F. Ben Nasr. Ensembles aléatoires self-affines en loi. *Bull. Sci. Math.*, **116** (1992), 111–119.

[11] J. D. Biggins. Martingale convergence in the branching random walk. *J. Appl. Probab.*, **14** (1977), 25–37.

[12] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Mathematics, no. 470. Springer-Verlag, Berlin-New York, 1975.

[13] T. Das, D. Simmons. The Hausdorff and dynamical dimensions of self-affine sponges: a dimension gap result. *Invent. Math.*, **210** (2017), 85–134.

[14] F. M. Dekking. On the survival probability of a branching process in a finite state i.i.d. environment. *Stochastic Process. Appl.* **27** (1987), 151–157.

[15] F. M. Dekking, G. R. Grimmett. Superbranching processes and projections of random Cantor sets. *Probab. Theory Related Fields*, **78** (1988), 335–355.

[16] R. Durrett, T. Liggett. Fixed points of the smoothing transformation. *Z. Wahrsch. Verw. Gebiete*, **64** (1983), 275–301.

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

[18] K. J. Falconer. Dimensions and measures of quasi self-similar sets. *Proc. Amer. Math. Soc.*, **106** (1989), 543–554.

[19] K. J. Falconer. Projections of random Cantor sets. *J. Theoret. Probab.,* **2** (1989), 65–70.

[20] K. J. Falconer, X. Jin. Exact dimensionality and projections of random self-similar measures and sets. *J. Lond. Math. Soc.*, **90** (2014), 388–412.

[21] K. Falconer, T. Kempton. Planar self-affine sets with equal Hausdorff, box and affinity dimensions. *Ergod.Th. & Dynam. Sys.*, **38** (2018), 1369–1388.

[22] K. Falconer and J. Miao. Dimensions of self-affine fractals and multifractals generated by upper-triangular matrices. *Fractals*, **15** (2007), 289–299.

[23] A. H. Fan, K. S. Lau, H. Rao. Relationships between different dimensions of a measure. *Monatsh. Math.*, **135** (2002), 191–201.

[24] K. Fan. Minimax theorems. *Proc. Nat. Acad. Sci. U.S.A.*, **39** (1953), 42–47.

[25] D. J. Feng. Dimension of invariant measures for self-affine iterated function system. arXiv preprint arXiv:1901.01691.

[26] D. J. Feng, W. Huang. Variational principle for weighted topological pressure. *J. Math. Pures Appl.*, **106** (2016), 411–452

[27] D. Gatzouras, S. Lalley. Statistically self-affine sets: Hausdorff and box dimensions. *J. Theoret. Probab.*, **7** (1994), 437–468.

[28] D. Gatzouras, Y. Peres. Invariant measures of full dimensions for some expanding maps. *Ergod. Th. & Dynam. Sys.*, **17** (1997), 147–167.

[29] S. Graf, R. D. Mauldin, S. C. Williams. The exact Hausdorff dimension in random recursive constructions. *Mem. Amer. Math. Soc.*, **71** (1988), no. 381.

[30] M. Hochman. Dimension theory of self-similar sets and measures. Proceedings of the International Congress of MathematiciansRio de Janeiro 2018. Vol. III. Invited lectures, 19491972, World Sci. Publ., Hackensack, NJ, 2018.

[31] M. Hochman, A. Rapaport. Hausdorff dimension of planar self-affine sets and measures with overlaps. arXiv preprint arXiv:1904.09812.

[32] J. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, **30** (1981) 713–747.

[33] T. Jordan, M. Pollicott and K. Simon. Hausdorff dimension for randomly perturbed self affine attractors. *Comm. Math. Phys.*, **270** (2007), 519–544.

[34] A. Käenmäki. On natural invariant measures on generalised iterated function systems. *Ann. Acad. Sci. Fenn. Math.*, **29** (2004), 419–458.

[35] A. Käenmäki, M. Vilppolainen. Dimension and measures on sub-self-affine sets. *Monatsh. Math.*, **161** (2010), 271–293.

[36] J.-P. Kahane. Multiplications aléatoires et dimensions de Hausdorff. *Ann. Inst. H. Poincaré Probab. Statist.*, **23** (1987), 289–296.

[37] J.-P. Kahane and J. Peyrière, Sur certaines martingales de B. Mandelbrot. *Adv. Math.*, **22** (1976), 131–145.

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

[39] Y. Kifer. Fractal dimensions and random transformations. *Trans. Amer. Math. Soc.*, **348** (1996), 2003–2038.

[40] I. Kolossvary, K. Simon. Triangular Gatzouras-Lalley-type planar carpets with overlaps. arXiv preprint arXiv:1807.09717.

[41] S. Lalley, D. Gatzouras. Hausdorff and box dimensions of certain self-affine fractals. *Indiana Univ. Math. J.*, **41** (1992), 533–568.

[42] F. Ledrappier and P. Walters. A relativised variational principle for continuous transformations. *J. Lond. Math. Soc.*, **16** (1977), 568–576.

[43] B. B. Mandelbrot, Intermittent turbulence in self-similar cascades, divergence of high moments and dimension of the carrier. *J. Fluid. Mech.*, **62** (1974), 331–358.

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

[45] S. M. Ngai. A dimension result arising from the L^q-spectrum of a measure. *Proc. Amer. Math. Soc.*, **125** (1997), 2943–2951.

[46] J. Peyrière. Comparaison de deux notions de dimension. *Bull. Soc. Math. France*, **114** (1986), 97–103.

[47] M. Rams, K. Simon. The geometry of fractal percolation. In: Geometry and analysis of fractals, 303–323, Springer Proc. Math. Stat., 88, Springer, Heidelberg, 2014.

[48] B. Solomyak. Measure and dimension for some fractal families. *Math. Proc. Cambridge Philos. Soc.*, **124** (1998), 531–546.

[49] B. von Bahr, C. G. Esseen. Inequalities for the r-th absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Stat.*, **36** (1965), No. 1, 299–303.

[50] T. Watanabe. Exact Hausdorff measure on the boundary of a Galton-Watson tree. *Ann. Probab.*, **35** (2007), 1007–1038.

Laboratoire d'Analyse, Gomtrie et Applications, CNRS, UMR 7539, Université Sorbonne Paris Nord, CNRS, UMR 7539, F-93430, Villetaneuse, France

Email address: `barral@math.univ-paris13.fr`

Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong,

Email address: `djfeng@math.cuhk.edu.hk`