HAL open science

Dimensions of random statistically self-affine Sierpinski sponges in R^k

Julien Barral, De-Jun Feng

To cite this version:

Julien Barral, De-Jun Feng. Dimensions of random statistically self-affine Sierpinski sponges in R^k. 2020. hal-02458498v1

HAL Id: hal-02458498

https://hal.science/hal-02458498v1

Preprint submitted on 28 Jan 2020 (v1), last revised 5 Dec 2022 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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-similar measures supported on K. This formula presents a new feature with respect to the deterministic case or the random dynamical version. Then, we establish a variational principle expressing $\dim K$ as the supremum of the Hausdorff dimensions of statistically self-similar measures supported on K, which is shown to be uniquely reached. The value of $\dim K$ 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 obtained by Gatzouras and Lalley. This alternative concerns both the sharp lower and upper bounds for the dimension.

The value of the box counting dimension of K and its equality with $\dim K$ are also studied. We also obtain a variational formula for the Hausdorff dimensions of the natural orthogonal projections of K to the linear subspaces generated by the eigensubspaces of the diagonal endomorphism used to generate K (contrarily to what happens in the deterministic case, these projections are not of the same nature as K). Finally, we prove a dimension conservation formula associated to any Mandelbrot measure supported on K, that of its orthogonal projection to such subspace, and the dimension of almost every associated conditional measure.

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 fibers 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.

Key words and phrases. Mandelbrot measures, Hausdorff dimension, multifractals, phase transitions, large deviations, branching random walk in a random environment.

2010 *Mathematics Subject Classification*: 28A78, 28A80, 60F10, 60G42, 60G57, 60K40.

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 [29]). When X is an 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 [27] 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 $A \subset \{0,\dots,m_1-1\}\times\{0,\dots,m_2-1\}$ of cardinality at least 2; the Bedford-McMullen carpet K associated with A is the attractor of the system $S_A = \Big\{f_a : (x_1,x_2) \mapsto \Big(\frac{x_1+a_1}{m_1}, \frac{x_2+a_2}{m_2}\Big) : a = (a_1,a_2) \in A\Big\}$ of contractive affine maps of the Euclidean plane; note that by construction $K \subset [0,1]^2$. Set $N_i = \#\{a \in A : a_2 = i\}$ 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, 40] that

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

and

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

where dim 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, $\dim K = \dim_B K$ if and only if the positive N_i are all equal. Note that the possible dimension gap $\dim K < \dim_B K$ cannot hold for self-similar sets (see [16]). 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 [35] extends the results of [40] 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 the dynamical dimension of K coincide. Moreover, the dimension of any ergodic measure is given by the Ledrappier-Young formula

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

where $\Pi_i = (x_1,\dots,x_k) \mapsto (x_i,\dots,x_k)$. Also, in the variational principle, the supremum is uniquely reached 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 $\{M_i\}_{1\leq i\leq N}$ of linear automorphisms of $\mathbb{R}^k$ to itself whose norm subordinated to the Euclidean norm on $\mathbb{R}^k$ is smaller that $1/2$, it was shown [17, 32] that for $\mathcal{L}^{Nk}$-almost every choice of N translation vectors $v_1,\dots,v_N$ in $\mathbb{R}^k$, the Hausdorff dimension of the attractor K of the affine IFS $\{M_ix+v_i\}_{1\leq i\leq N}$ is the maximum of the Hausdorff dimensions of the natural projections of ergodic measures on $(\{1,\dots,N\}^{\mathbb{N}^*},\sigma)$ to K (here σ stands for the shift operation); in addition for such a measure μ, $\dim_H(\mu)=\min(k,\dim_L(\mu))$, where $\dim_L(\mu)$ is the so-called Lyapunov dimension of μ [17, 44, 31, 30]. A similar result holds for $\mathcal{L}^{Nk^2}$-almost every choice of contractive $\{M_i\}_{1\leq i\leq N}$ in some non empty open set, with a fixed $(v_1,\dots,v_N)$ [2]. In these contexts one has $\dim K=\dim_B K=\min(k,\dim_A K)$, 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 $\psi(1)\leq\log(m_2)$ and $\frac{\psi(1)}{\log(m_1)}+(\frac{1}{\log(m_2)}-\frac{1}{\log(m_1)})\log(m_2)$ otherwise, so that in this case the previous equality between dimensions only occurs exceptionally). If both $\{M_i\}_{1\leq i\leq N}$ and $(v_1,\dots,v_N)$ are fixed, the following stronger result appeared recently in the 2-dimensional case: if $\{M_i\}_{1\leq i\leq N}$ satisfies the strong irreducibility property and $\{M_i/\sqrt{|\det(M_i)|}\}_{1\leq i\leq N}$ generate a non-compact group in $GL_2(\mathbb{R})$, and if the IFS $\{f_i\}_{1\leq i\leq N}$ is exponentially separated, then $\dim K$ is the supremum of the Lyapounov dimensions of the self-affine measures supported on K (it is not known if this supremum is reached in general); here again for such a measure, Lyapunov and Hausdorff dimension coincide, and $\dim K=\dim_B K=\dim_A K$ [3, 28]. 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 [19]. There are also similar results in the case that the strong irreducibility fails but the M_i cannot be simultaneously reduced to diagonal automorphisms [20, 4, 3, 28].

Let us come back to self-affine carpets. Their study was further developed with the introduction of Gatzouras-Lalley carpets [38], with an application to the study of some non-conformal nonlinear repellers [25] 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 [32, 5] (see also [37] 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 [12] 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 exists 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 $\mathcal{L}^d$ [30]. 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 $\{0,\dots,m_1-1\}\times\{0,\dots,m_2-1\}$ 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 an box-counting dimensions of such sets and their higher dimensional versions have been determined in [36] (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 [24], and their self-similar versions have been investigated extensively (see e.g. [26, 46, 43]). 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 $K \neq \emptyset$,

$$\dim K = \frac{\psi(\alpha)}{\log(m_2)} \text{ where } \alpha = \max\Big(t, \frac{\log(m_2)}{\log(m_1)}\Big), \tag{1.2}$$

and

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

Moreover, $\dim K = \dim_B K$ iff $t = 1$ or all the positive $\mathbb{E}(N_i)$ are equal (we note that the value of $\dim K$ was previously obtained in [42, 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 [24] 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 $\dim K$, the authors of [24] adapt the approach used by Bedford in the deterministic case: Π_2 still denoting the orthogonal projection on the x_2-axis the lower bound for $\dim K$ is obtained by taking the maximum of $\dim \Pi_2(K)$ and the maximum of the lower bound for the Hausdorff dimension of certain random subsets of K. Each such subset E is obtained by considering the union of almost all the fibers $\pi_2^{-1}(\{(0, x_2)\}$ 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 fibers are controlled from bellow. This yields a lower bound for $\dim E$ thanks to a theorem of Marstrand. The upper bound for $\dim K$ 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-dimensional fibers mentioned above is obtained thanks to statistically self-similar branching measures in random environment, the dimension of which is relatively direct to get, and yields the dimension of the fiber. 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 $2 \le i \le k$. Then for $i \ge 3$ we would meet the much harder problem to estimate the Hausdorff dimension of fibers 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 $\dim \Pi_i(K)$ as well, a question that we will naturally consider. Also, for the upper bound, extending to higher dimensions the combinatorial argument used in [24] to get effective coverings seems impossible. However, we will see that to getting the box-counting dimension of K in the higher dimension cases is rather direct from the two dimensional case treated in [24].

We will develop a dimension theory for statistically self-affine Sierpinski sponges in $\mathbb{R}^k$, for any $k \ge 2$, 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),

$$\begin{aligned}\dim(\mu) &= \frac{1}{\log(m_1)}\dim_e(\mu) + \sum_{i=2}^{k}\Big(\frac{1}{\log(m_i)} - \frac{1}{\log(m_{-1})}\Big)\dim_e(\mu\circ\Pi_i^{-1}) \\ &= \frac{1}{\log(m_1)}\dim_e(\mu) + \Big(\frac{1}{\log(m_i)} - \frac{1}{\log(m_{-1})}\Big)\cdot\min\big(\dim_e(\mu), h_{\nu_i}(\Pi_i\circ f)\big), \qquad (1.3)\end{aligned}$$

where $\dim_e(\mu)$ is the dimension entropy 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 4.3); 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, 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 $\dim K$ proves that this maximal Hausdorff dimension of a Mandelbrot measure supported on K yields $\dim K$. 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 $\dim K$ (see Theorem 4.8); when $k = 2$, this does not reduce back to the argument developed in [24]. As a by product, we get an alternative to the proof by Kenyon and Peres [34] 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 2.5. Another difference between the deterministic or random dynamical Sierpinski sponges, and the random Sierpinski sponges studied in this paper is that the first two classes of objects

are stable under the natural projections Π_i, $2 \leq i \leq k$, while this is not the case for the third one. Our approach will also provide $\dim \Pi_i(K)$ (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 [14, 17]). 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 $\mathbb{N}$, $\mathbb{R}_+$ and $\mathbb{R}_+^*$ stand for the sets of non negative integer, non positive real numbers, and strictly positive real numbers respectively.

Let $k \geq 2$ be an integer. Assume that (X_i, T_i) $(i = 1, \ldots, k)$ are fulshifts 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 $i = 1, \ldots, k-1$ (this meaning that $\pi_i\left((x_n)_{n=1}^\infty\right) = (\pi(x_n))_{n=1}^\infty$ 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 $(\pi_i(x_1 \cdots x_n) = \pi_i(x_1) \cdots \pi_i(x_n))$.

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.

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

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

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, 40, 35, 6].

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

$$K_n = \{x \in X_1 : c_{x_i}(x_{|i-1}) = 1 \ \forall 1 \le i \le n\} = \bigcup_{u\in\mathcal{A}_1^n:\ \prod_{i=1}^n c_{u_i}(u_{|i-1})=1} [u]. \tag{2.2}$$

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

$$K = \bigcap_{n\ge1} K_n,$$

is the boundary of a non-degenerate Galton Watson tree with offspring distribution 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:

$$K = \bigcup_{a\in A} a\cdot K^a, \quad K^a = \bigcap_{n\ge0} \bigcup_{u\in\mathcal{A}_1^n:\ \prod_{i=1}^n c_{u_i}(au_{|i-1})=1} [u],$$

and we will call it 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 $\dim K$. 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

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

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

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

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 $(Y_n(u), \sigma(W_a(uv) : a \in \mathcal{A}_1, v \in \mathcal{A}_1^*))_{n\ge1}$ 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 [39], and is called Mandelbrot measure. The support of μ is the set

$$K_\mu = \bigcap_{n\ge0} \bigcup_{u\in\mathcal{A}_1^n:\ Q(u)>0} [u] \subset K,$$

where the inclusion $K_\mu \subset K$ follows from the assumption $\{W_a > 0\} \subset \{c_a(u) = 1\}$ for all $u \in \mathcal{A}_1^*$ and all $a \in \mathcal{A}_1$. Moreover, if $T'(1-) > 0$, then $(Y_n(u))_{n\ge1}$ is uniformly integrable (see [34, 11, 15]), and in this case note that 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. If $T'(1-) \leq 0$, either $\mu = 0$ almost surely (one says that μ is degenerate), or $\mathbb{P}(\exists\, a \in A,\, W_a = 1$ and $W_{a'} = 0$ if $a' \neq a) = 1$ (see (see [34, 11, 15] as well), and in this later case $T'(1-) = 0$, K_μ is a singleton and μ is a Dirac measure [15].

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 all Borel subsets of 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 positive and finite Borel measure ν on X_i, its entropy dimension is defined as

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

whenever the limit exists. If ν is T_i-invariant, one has $\dim_e(\nu) = h_\nu(T_i)$, the entropy of ν relative to T_i.

Due to results by Kahane and Peyrière [34, 33], 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 [21] and $\dim_e(\mu) = T'(1)$. In particular,

$$\dim_e(\mu) \leq \log \mathbb{E}(\#A). \tag{2.3}$$

Before stating our first result, we present a result deduced from [7] about the entropy dimension of $\Pi_{i*}\mu$ (in [7] we consider 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 $d_{\vec{\gamma}}$ with $\vec{\gamma} = ((\log \#\mathcal{A}_1)^{-1})$; but projecting from X_1 to X_i and considering the entropy dimension does 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*

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

This result will follow from the stronger fact that the L^q-spectrum of μ is differentiable at 1.

2.3. **The Hausdorff and box counting dimensions of K.** To state our result on $\dim K$, we need to recall some elements of the weighted thermodynamic formalism.

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

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

Then set $\widetilde{\mathcal{A}}_i = \{b \in \mathcal{A}_i : \mathbb{E}(N_b^{(i)}) > 0\}$. Without loss of generality we suppose that $\#\widetilde{\mathcal{A}}_k \ge 2$. 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 $\#\widetilde{\mathcal{A}}_i \ge 2$ for all $2 \le i \le k$.

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

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

where

$$\Pi_{i,j} = \pi_{j-1} \circ \cdots \circ \pi_i \quad \text{if } j > i \tag{2.4}$$

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

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

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

$$\frac{\mathrm{d}P^{\vec{\boldsymbol{\beta}}}(\cdot\phi, T_i))}{\mathrm{d}\theta}\Big|_\theta = \int_{\widetilde{X}_i} \phi(x) \,\mathrm{d}\nu_{\theta\phi}(x) = \nu_{\theta\phi}(\phi). \tag{2.5}$$

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

$$\phi_i(x) = (\gamma_1 + \cdots + \gamma_i) \log \mathbb{E}(N_{x_1}^{(i)}). \tag{2.6}$$

For this potential and $\vec{\boldsymbol{\beta}} = \vec{\boldsymbol{\gamma}}^i$, setting

$$P_i = P^{\vec{\boldsymbol{\gamma}}^i}(\cdot\phi_i, T_i),$$

(2.5) yields

$$P_i'(\theta) = (\gamma_1 + \cdots + \gamma_i) \sum_{b \in \mathcal{A}_i} \nu_{\theta\phi_i}([b]) \log \mathbb{E}(N_b^{(i)}) \quad (\theta \in \mathbb{R}). \tag{2.7}$$

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 the last section of the paper, namely Section 8.

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

$$\widetilde{\theta}_i = \frac{\gamma_1 + \cdots + \gamma_{i-1}}{\gamma_1 + \cdots + \gamma_i} \quad \text{if } 2 \le i \le k, \tag{2.8}$$

and define

$$\widetilde{I} = \left\{ i \in I : \exists\, \theta \in [\widetilde{\theta}_i, 1] \text{ such that } P_i'(\theta) \ge 0 \right\}$$

and $i_0 = \min(\widetilde{I})$, where by convention $\min(\emptyset) = k+1$. If $\widetilde{I} \neq \emptyset$, set

$$\theta_{i_0} = \min \left\{ \theta \in [\widetilde{\theta}_{i_0}, 1] : P_{i_0}'(\theta) \ge 0 \right\}.$$

If $\widetilde{I} = \emptyset$, set $\theta_{k+1} = 1$ and $\phi_{k+1} = \phi_k$.

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

$$\begin{aligned} \dim K &= \sup\{\dim_e^{\vec{\gamma}}(\mu) : \mu \text{ is a positive Mandelbrot measure supported on } K_\mu \subset K\} \\ &= \begin{cases} P_{i_0}(\theta_{i_0}) & \text{if } i_0 \le k \\ P_k(1) & \text{if } i_0 = k+1 \end{cases} \\ &= \inf \left\{ P_i(\theta) : i \in I,\, \widetilde{\theta}_i \le \theta \le 1 \right\}, \end{aligned} \tag{2.9}$$

and the supremum is uniquely attained at $\mu_{\nu_{\theta_{i_0}} \phi_{i_0}}$.

Remark 2.4. *We used an abuse of notation. Indeed, in Theorem 2.3 the supremum must be understood as taken over the joint law of (K, μ) with μ a non-degenerate Mandelbrot measure supported on K.*

Theorem 2.5 (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 the necessary and sufficient condition for $\dim K = \dim_B K$. Define

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

For each $2 \le i \le k$, 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 2.7 about the necessary and sufficient condition for the equality $\dim K = \dim_B K$ to hold.

Lemma 2.6. *Each ψ_i takes the value $\log \mathbb{E}(\#A)$ at $\theta = 1$. Moreover, if $2 \le i \le k-1$, then $\psi_i \ge \psi_{i+1}$, and $\widehat{\theta}_i < 1$ implies $\widehat{\theta}_{i+1} < 1$.*

Proof. The first property is due to the relation $\mathbb{E}(N_{\widetilde{b}}^{(i)}) = \sum_{b \in \widetilde{\mathcal{A}}_i : \pi_i(b) = \widetilde{b}} \mathbb{E}(N_b^{(i)})$ for any $\widetilde{b} \in \widetilde{\mathcal{A}}_{i+1}$, and the second one is due both to this property and the subadditivity of $y \ge 0 \mapsto y^\theta$. The third property is a direct consequence of the two first ones. $\square$

Corollary 2.7. *It holds that* $\dim K = \dim_B K$ *with probability 1, conditional on* $K \neq \emptyset$*, if and only if* $\mathbb{E}(N_b^{(i)})$ *does not depend on* $b \in \widetilde{\mathcal{A}}_i$ *for all* $i \in I$ *such that* $\widehat{\theta}_i < 1$*.*

Next we present our results regarding the images of μ and K under the projections Π_i, $2 \leq i \leq k$.

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 a non-denenerate Mandelbrot measure such that $K_\mu \subset K$ almost surely.

Theorem 2.8 (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*

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

Now define $\overline{\theta}_i = 0$ and $\overline{\theta}_j = \widetilde{\theta}_j$ if $i < j \leq k$ (recall (2.8)). Set $I_i = \{i, \ldots, k\}$ and define $\widetilde{I}_i = \{j \in I_i : \exists \theta \in [\overline{\theta}_j, 1],\ P_j'(\theta) \geq 0\}$. Then define $j_0 = \min(\widetilde{I}_i)$ if $I_i \neq \emptyset$ and $j_0 = k+1$ otherwise. Also, set

$$\theta_{j_0} = \min\{\theta \in [\overline{\theta}_{j_0}, 1] : P_{j_0}'(\theta) \geq 0\} \text{ if } j_0 \leq k.$$

If $\widetilde{I}_i = \emptyset$, set $\theta_{k+1} = 1$.

Theorem 2.9 (Hausdorff dimension of $\Pi_i(K)$)**.** *Let* $2 \leq i \leq k$*. With probability 1, conditional on* $K \neq \emptyset$*,*

$$\dim \Pi_i(K) = \sup\{\dim_e^{\vec{\gamma}^i}(\Pi_{i*}\mu) : \mu \textit{ is a positive Mandelbrot measure supported on } K_\mu \subset K\}$$

$$= \begin{cases} P_{j_0}(\theta_{j_0}) & \textit{if } j_0 \leq k \\ P_{j_0}(1) & \textit{if } j_0 = k+1 \end{cases}, \tag{2.10}$$

and the supremum is uniquely attained if and only if $j_0 > i$*, or if* $j_0 = i$ *and (*$\theta_{j_0} > 0$ *or* $\theta_{j_0} = 0$ *and* $P_{j_0}'(0) = 0$*). In any of these cases the supremum is reached at* $\mu_{\nu_{\theta_{j_0}\phi_{j_0}}}$*.*

Remark 2.10. *We deduce from Theorems 2.3 and 2.9 that: (i) if* $2 \leq i < i_0$ *then* $\dim \Pi_i(K) = \dim K$*. (ii)* $\dim K = \dim \Pi_{i_0}(K)$ *if and only if* $i_0 \leq k$ *and* $P_{i_0}'(\theta_{i_0}) = 0$*, or if* $i_0 = k+1$*. (iii) If* $2 \leq i < j \leq k$ *then (a) if* $2 \leq i < j < j_0$*, then* $\dim \Pi_j(K) = \dim \Pi_i(K)$*; (b)* $\dim \Pi_{j_0}(K) = \dim \Pi_i(K)$ *if and only if* $j_0 \leq k$ *and* $P_{j_0}'(\theta_{j_0}) = 0$*, or* $j_0 = k+1$*.*

Theorem 2.11 (Box counting dimension of $\Pi_i(K)$)**.** *Let* $2 \leq i \leq k$*. With probability 1, conditional on* $K \neq \emptyset$*,*

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

Corollary 2.12. *Let* $2 \leq i \leq k$*. It holds that* $\dim \Pi_i(K) = \dim_B \Pi_i(K)$ *with probability 1, conditional on* $K \neq \emptyset$*, if and only if either of the three following 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, if* j_0' *is the maximum of those* $j \in I_i$ *such that for all* $j' \leq j$ *in* I_i *either* $0 < \widehat{\theta}_{j'} < 1$ *or* $\widehat{\theta}_{j'} = 0$ *and* $\psi_{j'}(0) = 0$*: (i) for all* $j \in I_i$ *such that* $j \leq j_0'$*, 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* $\sum_{b' \in \Pi_{i,j}^{-1}(\{b\})} \mathbb{E}(N_{b'}^{(i)})^{\widehat{\theta}_i}$ *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, the last three results are new except in the case $i = k$, which is reducible to the two dimensional case which follows from [14, 17, 7].

2.5. **Dimensions of conditional measures.** Given a non-degenerate Mandelbrot measure μ, conditional on $\mu \neq 0$, for each $2 \leq i \leq k$, 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]:

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* $\mathbb{E}(\Pi_{i*}\mu)$*. 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.* $x \in K$*,* $\lim_{n\to\infty} \frac{\log(\mu^z([x_{|n}]))}{-n} = \dim_e(\mu) - \dim_e(\Pi_{i*}\mu)$*, i.e.* $\dim_e(\mu) - h_{\nu_i}(T_i)$ *if* $\dim_e(\mu) > h_{\nu_i}(T_i)$ *and 0 otherwise. In particular the entropy dimension of* μ^z *exists and* $\dim_e(\mu^z) = \dim_e(\mu) - \dim_e(\Pi_{i*}\mu)$.

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 [18] 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 2.14. *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*

$$\dim(\mu^z) = \gamma_1(\dim_e(\mu) - h_{\nu_i}(T_i)) + \sum_{j=2}^{i-1} \gamma_j \big(\min(\dim_e(\mu), h_{\nu_j}(T_j)) - h_{\nu_i}(T_i) \big). \tag{2.11}$$

(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, $2 \leq i \leq j \leq k$.

Theorem 2.15. *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*

$$\dim((\Pi_{i*}\mu)^{j,z})) = \sum_{j'=i}^{j-1} \vec{\gamma}^i_{j'}\big(\min(\dim_e(\mu), h_{\nu_{j'}}(T_{j'})) - h_{\nu_j}(T_j))\big). \tag{2.12}$$

(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 $\mathcal{G}_n = \{\prod_{i=1}^k [\ell_i m_i^{-n}, (\ell_i + 1)m_i^{-n}] : 0 \leq \ell_i \leq m_i^n - 1\}$, 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 $\widetilde{K}$ (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 [34] (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 [22]. If $k = 2$, the Euclidean version of Theorem 2.2 yields the value of $\dim \widetilde{K}$ computed by Gatzouras and Lalley in [24].

Regarding the box counting dimension of $\widetilde{K}$, if $\widetilde{K}$ is deterministic, we just recover the result of [34]; 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 [24].

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 2.5 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 notations of 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

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

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 centered 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 positive and 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 $\tau'_\nu(1^+) \leq \underline{\dim}_{\text{loc}}(\nu, x) \leq \overline{\dim}_{\text{loc}}(\nu, x) \leq \tau'_\nu(1^-)$, so that the existence of $\tau'_\nu(1)$ implies the exact dimensionality of ν, with dimension equal to $\tau'_\nu(1)$ (see, e. g., [41]). Consequently, Theorem 2.2 follows from the following stronger one.

Theorem 3.1. *Suppose that $T(q) > -\infty$ for some $q > 1$. Conditional on $\mu \neq 0$, $\tau'_\mu(1)$ exists and equals $\dim_e^{\vec{\gamma}}(\mu)$.*

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

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

The theorem 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*

$$\mathbb{E}\Big(\sum_{B\in\mathcal{F}_n}\mu(B)^q\Big) = O\Big(\exp\Big(\ell_{i_0}(n)(c_0(q-1)^2 - T(q)) - \sum_{i=i_0+1}^{k}(\ell_i(n) - \ell_{i-1}(n))\mathcal{T}_{\nu_i}(q)\Big)\Big) \tag{3.1}$$

as $n \to \infty$. 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 $2 \leq i \leq k$.

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

The proof of Proposition 3.2 requires the following two 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. □

Lemma 3.4. [45] *Let $(L_j)_{j\geq 1}$ be a sequence of centered independent real valued random variables. For every finite $I \subset \mathbb{N}$ and $q \in (1, 2]$,*

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

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

$$B(x, e^{-\frac{n}{\gamma_1}}) = \{y \in X_1 : \forall\, 1 \leq i \leq k,\ \Pi_i(T_1^{\ell_{i-1}(n)} y) \in [U_i]\} = \bigcup_{(J_1,\ldots,J_k)\in\mathcal{J}_U} [J_1 \cdots J_k],$$

where

$$\mathcal{J}_U := \left\{(J_1, \ldots, J_k) \in \prod_{i=1}^{k} \mathcal{A}_1^{\ell_i(n)-\ell_{i-1}(n)} :\ \forall\, 1 \leq i \leq k,\ \Pi_i(J_i) = U_i\right\}. \tag{3.2}$$

For $q \in \mathbb{R}_+$ we need to estimate from above the partition function

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

For $1 \le i \le k$, set $\mathcal{U}_n^{(i)} = \prod_{j=i}^k \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$, and for $U^{(i)} = (U_i, \cdots, U_k) \in \mathcal{U}_n^{(i)}$, set

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

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

Then, for $1 \le i \le k$ 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

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

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 $0 \le i \le k-1$ we have

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

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 $\mu^{J_1 \cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}(J_{i+1})$. 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 $q > 1$:

$$\begin{aligned}
&\mathbb{E}\Big(\sum_{\substack{J_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)} \\ \Pi_i(J_{i+1}) = U_{i+1}}} \mu^{J_1 \cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([J_{i+1}])\, S_{U^{(i+2)}}(J_1 \cdots J_i J_{i+1}) \Big)^q \\
&\le 2^{q-1} \mathbb{E}\Big(\sum_{\substack{J_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)} \\ \Pi_i(J_{i+1}) = U_{i+1}}} \mu^{J_1 \cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([J_{i+1}]) \Big)^q \mathbb{E}(S_{U^{(i+2)}})^q \\
&\quad + 2^{q-1} \mathbb{E}\Big| \sum_{\substack{J_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)} \\ \Pi_i(J_{i+1}) = U_{i+1}}} \mu^{J_1 \cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([J_{i+1}])\, L(J_{i+1}) \Big|^q.
\end{aligned}$$

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

$$\begin{aligned}
&\mathbb{E}\Big|\sum_{\substack{J_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}\\ \Pi_i(J_{i+1})=U_{i+1}}} \mu^{J_1\cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([J_{i+1}])\, L(J_{i+1})\Big|^q\\
&\le 2^q\mathbb{E}\Big(\sum_{\substack{J_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}\\ \Pi_i(J_{i+1})=U_{i+1}}} \mu^{J_1\cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([J_{i+1}])^q\Big)\,\mathbb{E}(|L|^q)\\
&\le 2^q\mathbb{E}\Big(\sum_{\substack{J_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}\\ \Pi_i(J_{i+1})=U_{i+1}}} \mu^{J_1\cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([J_{i+1}])\Big)^q\,\mathbb{E}(|L|^q) \quad \text{(using superadditivity)}\\
&= 2^q\mathbb{E}\Big(\Pi_{i*}\mu^{J_1\cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([U_{i+1}])^q\Big)\,\mathbb{E}(|L|^q),
\end{aligned}$$

where $\mathbb{E}(|L|^q) = \mathbb{E}(|S_{U^{(i+2)}} - \mathbb{E}(S_{U^{(i+2)}})|^q) \le 2^q\mathbb{E}(S^q_{U^{(i+2)}})$, and $\mathbb{E}(S^q_{U^{(i+2)}})$ 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

$$\begin{aligned}
&\mathbb{E}\Big(\sum_{\substack{J_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}\\ \Pi_i(J_{i+1})=U_{i+1}}} \mu^{J_1\cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}(J_{i+1})\, S_{U^{(i+2)}}(J_1\cdots J_{i+1})\Big)^q\\
&\quad\le 2^{3q}\mathbb{E}\Big(\Pi_{i*}\mu^{J_1\cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([U_{i+1}])^q\Big)\mathbb{E}(S^q_{U^{(i+2)}}).
\end{aligned}$$

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

$$\begin{aligned}
&\mathbb{E}(Z_{q,n}(J_1\cdots J_i))\\
&\le 2^{3q}\sum_{U_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}} \mathbb{E}\Big(\Pi_{i*}\mu^{J_1\cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([U_{i+1}])^q\Big)\sum_{U^{(i+2)}\in\mathcal{U}_n^{(i+2)}}\mathbb{E}(S_{U^{(i+2)}}(J_1\cdots J_i\widetilde{J}_{i+1})^q)\\
&= 2^{3q}\sum_{U_{i+1}\in\mathcal{A}_{i+1}^{\ell_{i+1}(n)-\ell_i(n)}} \mathbb{E}\Big(\Pi_{i*}\mu^{J_1\cdots J_i}_{\ell_{i+1}(n)-\ell_i(n)}([U_{i+1}])^q\Big)\,\mathbb{E}(Z_{q,n}(J_1\cdots J_i\widetilde{J}_{i+1})).
\end{aligned}$$

Since $(\mu_p)_{p\ge1}$ and $(\mu_p^{J_1\cdots J_i})_{p\ge1}$ are identically distributed this yields

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

It follows that

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

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

$$\begin{aligned}\mathbb{E}(Z_{q,n}) &\le 2^{3qk} C_q^{k-1} \mathbb{E}(Y^q)\, \mathbb{E}\Big(\sum_{U_1 \in \mathcal{A}_1^n} \mu_n([U_1])^q \Big) \\ &\qquad \times \ \exp\Big(- \sum_{i=2}^k (\ell_i(n) - \ell_{i-1}(n)) \min(T(q), \mathcal{T}_{\nu_i}(q)) \Big) \\ &= 2^{3qk} C_q^{k-1} \mathbb{E}(Y^q)\, \exp\Big(- nT(q) - \sum_{i=2}^k (\ell_i(n) - \ell_{i-1}(n)) \min(T(q), \mathcal{T}_{\nu_i}(q)) \Big).\end{aligned}$$

Finally, recall that $i_0 = \max\{2 \le i \le k : T'(1) \le \mathcal{T}'_{\nu_i}(1)\}$ (note that for each i the numbers $\mathcal{T}'_{\nu_i}(1)$ is the measure theoretic entropy of ν_i so that the sequence $(\mathcal{T}'_{\nu_i}(1))_{1 \le i \le k}$ is non increasing). Since T and the functions $\mathcal{T}_{\nu_i}$ are analytic near 1 and coincide at 1, for all $2 \le i \le i_0$ there exists $q_{0,i} \in (1, q_1]$ and $c_i \ge 0$ such that for all $q \in (1, q_{0,i}]$ one has $\min(T(q), \mathcal{T}_{\nu_i}(q)) \ge T(q) - c_i(q-1)^2$, 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 (3.1).

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

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

so

$$\begin{aligned}\mathbb{E}(Z_{q,n}(J_1 \cdots J_i)) &\le \sum_{U_{i+1} \in \mathcal{A}_{i+1}^{\ell_{i+1}(n) - \ell_i(n)}} \sum_{\substack{J_{i+1} \in \mathcal{A}_1^{\ell_{i+1}(n) - \ell_i(n)} \\ \Pi_i(J_{i+1}) = U_{i+1}}} \mathbb{E}(\mu^{J_1 \cdots J_i}_{\ell_{i+1}(n) - \ell_i(n)}(J_{i+1})^q) \\ &\qquad \cdot \mathbb{E}\Big(\sum_{U^{(i+2)} \in \mathcal{U}_n^{(i+2)}} S_{U^{(i+2)}}(J_1 \cdots J_i \widetilde{J}_{i+1})^q \Big) \\ &= \mathbb{E}\Big(\sum_{J_{i+1} \in \mathcal{A}_1^{\ell_{i+1}(n) - \ell_i(n)}} \mu^{J_1 \cdots J_i}_{\ell_{i+1}(n) - \ell_i(n)}(J_{i+1})^q \Big) \mathbb{E}(Z_{q,n}(J_1 \cdots J_i \widetilde{J}_{i+1})) \\ &= \mathbb{E}\Big(\sum_{J_{i+1} \in \mathcal{A}_1^{\ell_{i+1}(n) - \ell_i(n)}} \mu_{\ell_{i+1}(n) - \ell_i(n)}(J_{i+1})^q \Big) \mathbb{E}(Z_{q,n}(J_1 \cdots J_i \widetilde{J}_{i+1})) \\ &= \exp(-(\ell_{i+1}(n) - \ell_i(n))T(q)) \mathbb{E}(Z_{q,n}(J_1 \cdots J_i \widetilde{J}_{i+1})).\end{aligned}$$

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

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

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

$$\begin{aligned}
&Z_{q,n}(\widetilde{J}) \\
&= \sum_{U^{(i_0+1)} \in \mathcal{U}_n^{(i_0+1)}} \Big(\sum_{\substack{J' = J_{i_0+1} \cdots, J_{k-1} \\ \Pi_j(J_j) = U_j,\ \forall\, i_0+1 \le j \le k-1}} \mu^{\widetilde{J}}_{\lambda(n)}([J']) \sum_{J_k \colon \Pi_k(J_k) = U_k} \mu^{\widetilde{J}J'}([J_k]) \Big)^q \\
&= \sum_{U^{(i_0+1)} \in \mathcal{U}_n^{(i_0+1)}} \Big(\sum_{\substack{J' = J_{i_0+1} \cdots J_{k-1} \\ \Pi_j(J_j) = U_j,\ \forall\, i_0+1 \le j \le k-1}} \mu^{\widetilde{J}}_{\lambda(n)}([J']) \, \nu_k([U_k]) \, X(\widetilde{J}J') \Big)^q,
\end{aligned}$$

where

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

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

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

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

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

this yields

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

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 $\nu_{k-1} = \mathbb{E}(\Pi_{k-1*} \mu^{\widetilde{J}J_{i_0+1} \cdots J_{k-2}}_{\ell_{k-1}(n) - \ell_{k-2}(n)})$. Iterating we get

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

and finally

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

□

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 $\dim K$. To do so, it is convenient to first relate $\dim_e(\mu)$ to $h_{\nu_i}(T_i)$ for all $2 \le i \le k$. 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 $\dim K$. Finally, in Section 4.4 we develop a kind of variational principle to get the sharp upper bound for $\dim K$.

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

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

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

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

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

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

with the conventions $0^0 = 0$ and $\log(0) = -\infty$. One can check that

$$e^{-T_W(q)} = \sum_{\substack{b\in\mathcal{A}_i \\ \nu_i([b])>0}} \nu_i([b])^q e^{-T_{V_b^{(i)}}(q)}, \tag{4.1}$$

from what it follows, after differentiating at 1, that

$$\dim_e(\mu) = h_{\nu_i}(T_i) + \dim_e(\mu|\nu_i), \tag{4.2}$$

where

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

is the entropy of the invariant measure ν_i and

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

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, 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

$$\dim_e(\mu_{\nu_i}|\nu_i) = \sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \mathbb{E}(N_b^{(i)}). \tag{4.3}$$

Remark 4.1. *The following basic observation will play an important role. Given a non-degenerate Mandelbrot measure μ supported on K and $2 \leq i \leq k$, 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 $-\log \mathbb{E}(N_b^{(i)})$ at 0, so it 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 (4.2) that $T'_{\widetilde{W}}(1) > 0$, and μ_{ν_i} is non-degenerate.*

Remark 4.2. *The reader will also check that when μ_{ν_i} is non-degenerate, for all $2 \leq i' \leq i-1$, 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*

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

Also, if ν_i is fully supported on $\widetilde{X}_i$, then $K_{\mu_{\nu_i}} = K$ almost surely.

4.2. **An optimisation problem.** The following result invokes several definitions given in Section 1, especially in Sections 2.2 and 2.3.

Theorem 4.3. *Let*

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

On has

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

and the maximum is uniquely attained at $\mu_{\nu_{\theta_{i_0}}\phi_{i_0}}$.

Now let us introduce some definitions and make some observations (Remark 4.4).

For $1 \le i \le k$, set $\widetilde{\gamma}_i = \gamma_1 + \cdots + \gamma_i$.

Given a non-degenerate Mandelbrot measure μ supported on K, ν_j still standing for $\mathbb{E}(\Pi_{j*}\mu)$, for $2 \le i \le k$ define

$$\begin{aligned} d_i(\mu) &= \sum_{j=1}^{i-1} \gamma_j \dim_e(\mu) + \sum_{j=i}^{k} \gamma_j h_{\nu_j}(T_j) \\ &= \widetilde{\gamma}_{i-1} \left(\sum_{b \in \mathcal{A}_i} \nu_i([b])\, T'_{V_b^{(i)}}(1) \right) + \widetilde{\gamma}_i h_{\nu_i}(T_i) + \sum_{j=i+1}^{k} \gamma_j h_{\nu_j}(T_j) \\ \text{and } d_{k+1}(\mu) &= \widetilde{\gamma}_k \left(\sum_{b \in \mathcal{A}_k} \nu_k([b])\, T'_{V_b^{(k)}}(1) \right) + \widetilde{\gamma}_k h_{\nu_k}(T_k). \end{aligned}$$

For $2 \le i \le k$, let $\mathcal{B}_i$ be the set of Bernoulli product measures ν_i on X_i whose topological support is included in $\widetilde{X}_i$, i.e. such that $\nu_i([b]) = 0$ if $b \notin \widetilde{\mathcal{A}}_i$, and ν_i is not a Dirac mass. Also, for $\nu_i \in \mathcal{B}_i$ define

$$D_i(\nu_i) = \widetilde{\gamma}_{i-1} \left(\sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \mathbb{E} N_b^{(i)} \right) + \widetilde{\gamma}_i h_{\nu_i}(T_i) + \sum_{j=i+1}^{k} \gamma_j h_{\nu_j}(T_j) \tag{4.4}$$

$$\text{and } D_{k+1}(\nu_k) = \widetilde{\gamma}_k \left(\sum_{b \in \mathcal{A}_k} \nu_k([b]) \log \mathbb{E} N_b^{(k)} \right) + \widetilde{\gamma}_k h_{\nu_k}(T_k), \tag{4.5}$$

where ν_j stands for the projection of ν_i to X_j.

Given $\nu_i \in \mathcal{B}_i$ and the random vector $\widetilde{W} = (\nu_i(b)\widetilde{V}_{b,a}^{(i)})_{a \in \mathcal{A}_1, b = \Pi_i(a)}$ defined in Section 4.1, if $T'_{\widetilde{W}}(1) > 0$, recall that the associated non-degenerate Mandelbrot measure μ_{ν_i} is so that ν_i coincides with $\mathbb{E}(\Pi_{i*}\mu_{\nu_i})$. We also define $\nu_{i'}$ as $\mathbb{E}(\Pi_{i'*}\mu)$ for $1 \le i' < i$. We also note that if $\sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \mathbb{E} N_b^{(i)} \ge 0$, it is direct from the differentiation of (4.1) at 1 that $T'_{\widetilde{W}}(1) > 0$ since ν_i is not a Dirac mass; moreover, $d_i(\mu_{\nu_i}) = D_i(\nu_i)$.

The following subsets of the $\mathcal{B}_i$ will play a natural role. Let

$$\begin{aligned} \widetilde{\mathcal{B}}_i &= \left\{ \nu_i \in \mathcal{B}_i : \begin{cases} \displaystyle\sum_{b \in \mathcal{A}_{i-1}} \nu_{i-1}([b]) \log \mathbb{E} N_b^{(i-1)} \le 0 \\ \sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \mathbb{E} N_b^{(i)} \ge 0 \end{cases} \right\} \quad (2 \le i \le k) \\ \text{and } \widetilde{\mathcal{B}}_{k+1} &= \Big\{ \nu_k \in \mathcal{B}_k : \sum_{b \in \mathcal{A}_k} \nu_k([b]) \log \mathbb{E} N_b^{(k)} \le 0 \Big\}. \end{aligned}$$

Then we set

$$D_i = \max\{D_i(\nu_i) : \nu_i \in \widetilde{\mathcal{B}}_i\} \quad (2 \le i \le k) \tag{4.6}$$

$$\text{and } D_{k+1} = \max\{D_{k+1}(\nu_k) : \nu_k \in \widetilde{\mathcal{B}}_{k+1}\}, \tag{4.7}$$

with the convention that $\max(\emptyset) = -\infty$.

Remark 4.4. *Let $2 \le i \le k$, $\nu_i \in \mathcal{B}_i$, and $\nu_1, \ldots, \nu_{i-1}$ defined from the non-degenerate Mandelbrot measure μ_{ν_i} as above.*

(1) If $i=2$, the inequality $\sum_{b\in\mathcal{A}_1}\nu_1([b])\log\mathbb{E}N_b^{(1)}\le 0$ always holds and it can be strict only if K is not deterministic since $N_b^{(i)}\le 1$ for all $b\in\mathcal{A}_1$ almost surely.

Also, due to Remark 4.2 and the fact that the entropy cannot increase after projection, if $i\ge 3$, and the measure μ_{ν_i} is non-degenerate, then $\sum_{b\in\mathcal{A}_i}\nu_i([b])\log\mathbb{E}N_b^{(i)}\le 0$ implies $\sum_{b'\in\mathcal{A}_{i'}}\nu_{i'}([b'])\log\mathbb{E}N_{b'}^{(i')}\le 0$ for all $2\le i'<i$.

(2) If $i\ge 3$ and the measure μ_{ν_i} is non-degenerate, then

$$\begin{aligned}D_{i-1}(\nu_{i-1})-D_i(\nu_i)&=d_{i-1}(\mu_{\nu_i})-d_i(\mu_{\nu_i})\\&=\gamma_{i-1}\Big(h_{\nu_{i-1}}(T_{i-1})-h_{\nu_i}(T_i)-\sum_{b\in\mathcal{A}_i}\nu_i([b])\log\mathbb{E}N_b^{(i)}\Big).\end{aligned}$$

Consequently, $\sum_{b\in\mathcal{A}_i}\nu_i([b])\log\mathbb{E}N_b^{(i)}<0$ implies $d_{i-1}(\mu_{\nu_i})>d_i(\mu_{\nu_i})$. Also, if $i=k$, then $d_k(\mu_{\nu_k})-d_{k+1}(\mu_{\nu_k})=-\gamma_k\sum_{b\in\mathcal{A}_k}\nu_k([b])\log\mathbb{E}N_b^{(k)}$. Hence $d_k(\mu_{\nu_k})\ge d_{k+1}(\mu_{\nu_k})$ if and only if $\sum_{b\in\mathcal{A}_k}\nu_k([b])\log\mathbb{E}N_b^{(k)}\le 0$.

Theorem 4.3 follows from the two lemmas established below.

Lemma 4.5. *One has $M^{\vec\gamma}=\widetilde{M}^{\vec\gamma}:=\max\{D_i:2\le i\le k+1\}$.*

Proof. We first remark that given a non-degenerate Mandelbrot measure μ supported on K, if $2\le i\le k$, $\nu_i=\mathbb{E}(\Pi_{i*}\mu)$ and μ_{ν_i} constructed as in Section 4.1, we have $K_\mu\subset K_{\mu_{\nu_i}}$ almost surely and $\dim_e(\mu|\nu_i)\le\dim_e(\mu_{\nu_i}|\nu_i)$ conditional on $\{\mu\ne 0\}$ (see (4.3) and Remark 4.1), hence $\dim_e(\mu)\le\dim_e(\mu_{\nu_i})$. Also, by definition, $\dim_e^{\vec\gamma}(\mu)=\min\{d_2(\mu),\dots,d_{k+1}(\mu)\}$, hence $\dim_e^{\vec\gamma}(\mu)=d_i(\mu)$, where $i=\min\{2\le j\le k:\dim_e(\mu)\ge h_{\nu_j}(T_j)\}$, with the convention $\min(\emptyset)=k+1$.

The previous observations together with Remark 4.4 imply that: (1) the supremum $M^{\vec\gamma}$ we seek for is reached for a measure μ of the form μ_{ν_i}, $2\le i\le k$; (2) Suppose that such a Mandelbrot measure is given, and $\dim_e^{\vec\gamma}(\mu)=d_j(\mu)$ for some $2\le j\le k+1$; it is necessary that $\nu_j\in\widetilde{\mathcal{B}}_j$. □

Recall that i_0 was defined just before the statement of Theorem 2.3.

Lemma 4.6. *$\widetilde{M}^{\vec\gamma}$ is reached at D_{i_0}. Moreover, D_{i_0} is uniquely reached at $\nu_{i_0}=\nu_{\theta_{i_0}\phi_{i_0}}$. Also, $M^{\vec\gamma}$ is uniquely reached at $\mu_{\nu_{i_0}}$, and it is equal to*

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

Proof. It is rather long and will consist in distinguishing three situations.

At first we notice that the observation made in the last paragraph of the previous proof shows that if $2\le i\le k$ and $\mathbb{E}(N_b^{(j)})<1$ for all $2\le j<i$ and $b\in\mathcal{A}_j$, then $M^{\vec\gamma}\in\{D_i,\dots,D_{k+1}\}$ (recall (4.6) and (4.7)).

Now, let j_0 be the infinimum of the set of those $2\le j\le k$ such that $\mathbb{E}(N_b^{(j)})\ge 1$ for some $b\in\mathcal{A}_j$ if this set is not empty, and $j_0=k+1$ otherwise. Below we discuss the three situations $j_0=k+1$, $j_0=k$, and $j_0\le k-1$. The two first ones are enough to cover the case $k=2$.

• **Suppose** $j_0 = k+1$**.** It is necessary that $M^{\vec{\gamma}} = D_{k+1}$, and optimising yields

$$\nu_k([b]) = \frac{\mathbb{E}(N_b^{(k)})}{\sum_{b' \in \mathcal{A}_k} \mathbb{E}(N_{b'}^{(k)})}$$

for all $b \in \mathcal{A}_k$ and $M^{\vec{\gamma}} = D_{k+1}(\nu_k) = P^{(\widetilde{\gamma}_k)}(\phi_k, T_k) = P_k(1)$, that is $\nu_k = \nu_{\phi_k}$. Thus, $M^{\vec{\gamma}}$ is uniquely reached at the Mandelbrot measure μ_{ν_k} with $\nu_k = \nu_{\phi_k}$. Moreover $i_0 = k+1$.

• **Suppose** $j_0 = k$**.** We first consider the optimisation of $D_k(\nu_k)$ over $\nu_k \in \mathcal{B}_k$ rather than $\widetilde{\mathcal{B}}_k$.

Recall from Section 2.3 the definition of $P_k(\theta)$ and $\nu_{\theta\phi_k}$ for $\theta \in \mathbb{R}$. It is standard that

$$P_k(\theta) = \widetilde{\gamma}_k \log \sum_{b \in \mathcal{A}_k} \mathbb{E}(N_b^{(k)})^\theta \tag{4.8}$$

and $\nu_{\theta\phi_k}$ is the Bernoulli product measure such that

$$\nu_{\theta\phi_k}([b]) = \frac{\mathbb{E}(N_b^{(k)})^\theta}{\sum_{b' \in \mathcal{A}_k} \mathbb{E}(N_{b'}^{(k)})^\theta} \quad (\forall\, b \in \widetilde{\mathcal{A}}_i).$$

We also know that

$$P_k'(\theta) = \widetilde{\gamma}_k \sum_{b \in A_k} \nu_{\theta\phi_k}([b]) \log \mathbb{E}(N_b^k). \tag{4.9}$$

Now, recall that for $\nu_k \in \mathcal{B}_k$, $D_k(\nu_k)$ (see the definition (4.4)) rewrites

$$D_k(\nu_k) = \widetilde{\gamma}_k h_{\nu_k}(T_k) + \sum_{b \in \mathcal{A}_k} \nu_k([b]) \widetilde{\gamma}_k \frac{\widetilde{\gamma}_{k-1}}{\widetilde{\gamma}_k} \log \mathbb{E}(N_b^{(k)}) = \widetilde{\gamma}_k \Big(h_{\nu_k}(T_k) + \int_{\widetilde{X}_k} \widetilde{\theta}_k \phi_k(x)\, \mathrm{d}\nu_k(x) \Big),$$

where we recall that for $2 \le i \le k$, we defined

$$\widetilde{\theta}_i = \frac{\gamma_1 + \cdots + \gamma_{i-1}}{\gamma_1 + \cdots + \gamma_i}.$$

Consequently, the optimum of $D_k(\nu_k)$ equals $P_k(\widetilde{\theta}_k)$ and is uniquely reached at $\nu_k = \nu_{\widetilde{\theta}_k \phi_k}$. For simplicity we denote $\nu_{\widetilde{\theta}_k \phi_k}$ by $\widetilde{\nu}_k$.

We now distinguish two cases.

First case: $\sum_{b \in \mathcal{A}_k} \widetilde{\nu}_k([b]) \log \mathbb{E}(N_b^{(k)}) \ge 0$. In this case $\widetilde{\nu}_k \in \widetilde{B}_k$ and $D_k(\widetilde{\nu}_k) = D_k$. Moreover, if $\nu_k \in \mathcal{B}_k$ and $\sum_{b \in \mathcal{A}_k} \nu_k([b]) \log \mathbb{E}(N_b^{(k)}) \le 0$, i.e. $\nu_k \in \widetilde{B}_{k+1}$, then

$$D_{k+1}(\nu_k) = \widetilde{\gamma}_k h_{\nu_k}(T_k) + \widetilde{\gamma}_k \sum_{b \in \mathcal{A}_k} \nu_k([b]) \log \mathbb{E}(N_b^{(k)}) \le D_k(\nu_k),$$

with equality only if $\sum_{b \in \mathcal{A}_k} \nu_k([b]) \log \mathbb{E}(N_b^{(k)}) = 0$. Consequently $D_{k+1}(\nu_k) = D_k(\nu_k) = D_k(\widetilde{\nu}_k)$, only if $\nu_k = \widetilde{\nu}_k$. This implies that $M^{\vec{\gamma}}$ is uniquely reached at the Mandelbrot measure μ_{ν_k} with $\nu_k = \widetilde{\nu}_k$. Also, $i_0 = k$, and $\theta_{i_0} = \widetilde{\theta}_k$.

Second case: $\sum_{b \in \mathcal{A}_k} \widetilde{\nu}_k([b]) \log \mathbb{E}(N_b^{(k)}) < 0$. In this case, the maximum of $D_k(\nu_k)$ whenever ν_k describes $\widetilde{B}_k$ is reached under the constraint $\sum_{b \in \mathcal{A}_k} \nu_k([b]) \log \mathbb{E}(N_b^{(k)}) = 0$. Indeed, let $\widehat{\nu}_k \in \widetilde{B}_k$ at which this maximum is reached, and let $\widetilde{p}$ and $\widehat{p}$ be the probability vectors associated with $\widetilde{\nu}_k$ and $\widehat{\nu}_k$ respectively. If $\sum_{b \in \mathcal{A}_k} \widehat{\nu}_k([b]) \log \mathbb{E}(N_b^{(k)}) > 0$, let λ be the unique element of $(0,1)$ such that $\sum_{b \in \mathcal{A}_k} (\lambda \widetilde{\nu}_k([b]) + (1-\lambda)\widehat{\nu}_k([b])) \log \mathbb{E}(N_b^{(k)}) = 0$.

Then the element ν_k of $\widetilde{B}_k$ associated with $\lambda\widetilde{p}+(1-\lambda)\widehat{p}$ is such that $h_{\nu_k}(T_k) > \lambda h_{\widetilde{\nu}_k}(T_k) + (1-\lambda)h_{\widehat{\nu}_k}(T_k)$, so $D_k(\nu_k) > D_k(\widehat{\nu}_k)$. This contradicts the definition of $\widehat{\nu}_k$.

Now, we notice that our assumption on $\widetilde{\nu}_k$, namely $\sum_{b\in\mathcal{A}_k}\widetilde{\nu}_k([b])\log\mathbb{E}(N_b^{(k)}) < 0$, implies the strict convexity of P_k (see (4.8)). Then, using Lagrange multipliers, we see that our optimisation problem has a unique solution $\nu_k = \nu_{\theta\phi_k}$ in the case that there exists $\theta\in\mathbb{R}$ (necessarily unique) such that

$$\sum_{b\in\mathcal{A}_k}\nu_{\theta\phi_k}([b])\log\mathbb{E}(N_b^{(k)}) = 0.$$

Moreover, due to (4.9) and our assumption on $\widetilde{\nu}_k$ we have $P_k'(\widetilde{\theta}_k) < 0$. Thus, since P_k is convex, if such a $\theta = \theta_0$ exists we must have $\theta_0 \geq \widetilde{\theta}_k$. Also $\nu_{\theta_0\phi_k} \in \widetilde{B}_{k+1}$.

Now suppose that θ_0 does exist and recall that $\max\{D_{k+1}(\nu_k) : \nu_k \in \mathcal{B}_k\}$ is reached at ν_{ϕ_k}.

If $\theta_0 \leq 1$, by strict convexity of P_k we must have $P_k'(1) = \sum_{b\in\mathcal{A}_k}\nu_{\phi_k}([b])\log\mathbb{E}(N_b^{(k)}) \geq 0$, with equality if and only if $\theta_0 = 1$. If the inequality is strict, an argument already used above shows that D_{k+1} must be reached by a measure $\nu_k \in \widetilde{\mathcal{B}}_{k+1}$ such that $\sum_{b\in\mathcal{A}_k}\nu_k([b])\log\mathbb{E}(N_b^{(k)}) = 0$, for which $D_k(\nu_k)$ and $D_{k+1}(\nu_k)$ coincide, so we find that D_{k+1} it is reached at $\nu_{\theta_0\phi_k}$ as well; in particular $D_k = D_{k+1}$. If $\theta_0 = 1$, for similar reasons we still have $D_k = D_{k+1}$, both reached at ν_{ϕ_k}. Consequently, $M^{\vec{\gamma}} = D_k = D_{k+1}$ and this supremum is uniquely reached at the Mandelbrot measure μ_{ν_k} with $\nu_k = \nu_{\theta_0\phi_k}$. Notice also that in any case, $i_0 = k$ and $\theta_{i_0} = \theta_0$.

If $\theta_0 > 1$, this time the strict convexity of P_k implies both $D_{k+1}(\nu_{\phi_k}) > D_{k+1}(\nu_{\theta_0\phi_k}) = D_k(\nu_{\theta_0\phi_k})$ and $\sum_{b\in\mathcal{A}_k}\nu_{\phi_k}([b])\log\mathbb{E}(N_b^{(k)}) < 0$. So $\nu_{\phi_k} \in \widetilde{\mathcal{B}}_{k+1}$ and $D_{k+1} > D_k$. It follows that $M^{\vec{\gamma}}$ is uniquely reached at the Mandelbrot measure μ_{ν_k} with $\nu_k = \nu_{\nu_{\phi_k}}$. Here $i_0 = k+1$.

If there is no $\theta \in \mathbb{R}$ such that $\sum_{b\in\mathcal{A}_k}\nu_{\theta\phi_k}([b])\log\mathbb{E}(N_b^{(k)}) = 0$, this implies that $\mathbb{E}(N_b^{(k)}) \leq 1$ for all $b \in \mathcal{A}_k$ and D_k is reached at ν_k such that for $b \in \mathcal{A}_k$, $\nu_k([b]) > 0$ implies $\mathbb{E}(N_b^{(k)}) = 1$, which after optimisation yields $\nu_k([b]) = 1/\#\{b \in \mathcal{A}_k : \mathbb{E}(N_b^{(k)}) = 1\}$. Thus, $D_k = D_k(\nu_k) = \widetilde{\gamma}_k \log\#\{b \in \mathcal{A}_k : \mathbb{E}(N_b^{(k)}) = 1\} < \widetilde{\gamma}_k \log\sum_{b\in\mathcal{A}_k}\mathbb{E}(N_b^{(k)}) = D_{k+1}(\nu_{\phi_k}) = D_{k+1}$. Here, the strict inequality comes from the fact that since we have $\sum_{b\in\mathcal{A}_k}\widetilde{\nu}_k([b])\log\mathbb{E}(N_b^{(k)}) < 0$, there is some $b \in \mathcal{A}_k$ such that $\mathbb{E}(N_b^{(k)}) < 1$, and the last equality holds since all the $\log\mathbb{E}(N_b^{(k)})$ are non positive, hence $\nu_{\phi_k} \in \widetilde{\mathcal{B}}_{k+1}$. Finally $M^{\vec{\gamma}}$ is uniquely reached at the Mandelbrot measure μ_{ν_k} with $\nu_k = \nu_{\phi_k}$, and here again $i_0 = k+1$.

Remark 4.7. *The previous discussion proves that the conclusion of the lemma holds true when $k = 2$.*

• **Suppose that $j_0 \leq k-1$.** This assumes $k \geq 3$. As in the previous case we first consider the optimisation of $D_{j_0}(\nu_{j_0})$ over $\nu_{j_0} \in \mathcal{B}_{j_0}$. To this end we write

$$D_{j_0}(\nu_{j_0}) = h_{\nu_{j_0+1}}^{(\widetilde{\gamma}_{j_0+1},\ldots,\gamma_k)}(T_{j_0+1}) + \widetilde{\gamma}_{j_0}h_{\nu_{j_0}}(T_{j_0}) + \widetilde{\gamma}_{j_0-1}\sum_{b\in\mathcal{A}_{j_0}}\nu_{j_0}([b])\log\mathbb{E}(N_b^{(j_0)})$$

in the form (recall the definition of $\widetilde{\theta}_{i_0}$ in (4.2)):

(4.10)

$$D_{j_0}(\nu_{j_0}) = h_{\nu_{j_0+1}}^{(\widetilde{\gamma}_{j_0+1},\dots,\gamma_k)}(T_{j_0+1}) + \sum_{\widehat{b}\in\mathcal{A}_{j_0+1}} \nu_{j_0+1}([\widehat{b}])\cdot\widetilde{\gamma}_{j_0} \sum_{\substack{b\in\mathcal{A}_{j_0}\\ \pi_{j_0}([b])=[\widehat{b}]}} -p(b|\widehat{b})\log p(b|\widehat{b}) + p(b|\widehat{b})\widetilde{\theta}_{j_0}\log\mathbb{E}(N_b^{(j_0)}),$$

where

$$p(b|\widehat{b}) = \begin{cases} \dfrac{\nu_{j_0}([b])}{\nu_{j_0+1}([\widehat{b}])} & \text{if } \nu_{j_0+1}([\widehat{b}])>0 \\ 0 & \text{otherwise.}\end{cases}$$

By definition of j_0, we do have $\sum_{b'\in\mathcal{A}_{j_0-1}}\nu_{j_0-1}([b'])\log\mathbb{E}N_{b'}^{(j_0-1)} \le 0$ for all $\nu_{j_0}\in\mathcal{B}_{j_0}$. As mentioned above, we first ignore the requirement $\sum_{b\in\mathcal{A}_{j_0}}\nu_{j_0}([b])\log\mathbb{E}N_b^{(j_0)}\ge 0$ which would hold if we directly optimised over $\widetilde{\mathcal{B}}_{j_0}$. Then, the above expression for $D_{j_0}(\nu_{j_0})$ implies that optimising given ν_{j_0+1} yields $p(b|\widehat{b}) = p_{\widetilde{\theta}_{j_0}}(b|\widehat{b})$, where

$$p_\theta(b|\widehat{b}) = \begin{cases} \dfrac{\mathbb{E}(N_b^{(j_0)})^\theta}{\sum\limits_{b'\in\mathcal{A}_{j_0}:\,\pi_{j_0}([b'])=[\widehat{b}]}\mathbb{E}(N_{b'}^{(j_0)})^\theta} & \text{if } \nu_{j_0+1}([\widehat{b}])>0, \\ 0 & \text{otherwise}\end{cases}. \tag{4.11}$$

Thus, given $\nu_{j_0+1}\in\mathcal{B}_{j_0+1}$, if for $\theta\in\mathbb{R}$ we define

$$\begin{aligned} D_{j_0}(\nu_{j_0+1},\theta) &= h_{\nu_{j_0+1}}^{(\widetilde{\gamma}_{j_0+1},\dots,\gamma_k)}(T_{j_0+1}) \\ &\quad + \sum_{\widehat{b}\in\mathcal{A}_{j_0+1}} \nu_{j_0+1}([\widehat{b}])\cdot\widetilde{\gamma}_{j_0} \sum_{\substack{b\in\mathcal{A}_{j_0}\\ \pi_{j_0}([b])=[\widehat{b}]}} -p_\theta(b|\widehat{b})\log p_\theta(b|\widehat{b}) + p_\theta(b|\widehat{b})\cdot\theta\log\mathbb{E}(N_b^{(j_0)}) \\ &= h_{\nu_{j_0+1}}^{(\widetilde{\gamma}_{j_0+1},\dots,\gamma_k)}(T_{j_0+1}) + \sum_{\widehat{b}\in\mathcal{A}_{j_0+1}} \nu_{j_0+1}([\widehat{b}])\cdot\widetilde{\gamma}_{j_0}\log \sum_{\substack{b\in\mathcal{A}_{j_0}\\ \pi_{j_0}([b])=[\widehat{b}]}} \mathbb{E}(N_b^{(j_0)})^\theta, \end{aligned}$$

then we have

$$\begin{aligned} &\max\{D_{j_0}(\nu_{j_0}) :\ \nu_{j_0}\in\mathcal{B}_{j_0},\ \pi_{j_0*}\nu_{j_0} = \nu_{j_0+1}\} \\ &= D_{j_0}(\nu_{j_0+1},\widetilde{\theta}_{j_0}) = h_{\nu_{j_0+1}}^{(\widetilde{\gamma}_{j_0+1},\dots,\gamma_k)}(T_{j_0+1}) + \sum_{\widehat{b}\in\mathcal{A}_{j_0+1}} \nu_{j_0+1}([\widehat{b}])\cdot\widetilde{\gamma}_{j_0}\log \sum_{\substack{b\in\mathcal{A}_{j_0}\\ \pi_{j_0}([b])=[\widehat{b}]}} \mathbb{E}(N_b^{(j_0)})^{\widetilde{\theta}_{j_0}}. \end{aligned}$$

Set $\widetilde{\nu}_{j_0} = \nu_{\widetilde{\theta}_{j_0}\phi_{j_0}}$. By definition of this measure, denoting $\pi_{j_0*}\widetilde{\nu}_{j_0}$ by $\widetilde{\nu}_{j_0+1}$ we have

$$D_{j_0}(\widetilde{\nu}_{j_0+1},\widetilde{\theta}_{j_0}) = \sup\{D_{j_0}(\nu_{j_0+1},\widetilde{\theta}_{j_0}) : \nu_{j_0+1}\in\mathcal{B}_{j_0+1}\} = P^{(\widetilde{\gamma}_{j_0},\gamma_{j_0+1},\dots,\gamma_k)}(\phi_{\widetilde{\theta}_{j_0}\phi_{j_0}}, T_{j_0}),$$

and the supremum is uniquely reached at $\widetilde{\nu}_{j_0+1}$.

As when $j_0 = k$, two cases must be distinguished.

First case: $\sum_{b\in\mathcal{A}_{j_0}}\widetilde{\nu}_{j_0}([b])\log\mathbb{E}(N_b^{(j_0)})\ge 0$. Then $\widetilde{\nu}_{j_0}\in\widetilde{\mathcal{B}}_{j_0}$ and $d_{j_0}(\mu_{\widetilde{\nu}_{j_0}}) = D_{j_0}(\widetilde{\nu}_{j_0}) = D_{j_0}$. Also, given $\nu_{j_0+1}\in\mathcal{B}_{j_0+1}$, recalling that ν_{j_0} stands for $\mathbb{E}(\Pi_{j_0*}\mu_{\nu_{j_0+1}})$, if the inequality

$\sum_{b\in\mathcal{A}_{j_0}} \nu_{j_0}([b])\log\mathbb{E}(N_b^{(j_0)}) \le 0$ holds, which is the case if $\nu_{j_0+1}\in\widetilde{\mathcal{B}}_{j_0+1}$, by definition of the convex function $D_{j_0}(\nu_{j_0+1},\cdot)$, this function has a non positive derivative at 1. Indeed,

$$\begin{aligned}
D'_{j_0}(\nu_{j_0+1},1) &= \sum_{\widehat{b}\in\mathcal{A}_{j_0+1}} \nu_{j_0+1}([\widehat{b}])\cdot\widetilde{\gamma}_{j_0}\frac{\sum_{b\in\mathcal{A}_{j_0},\,\pi_{j_0}([b])=[\widehat{b}]}\mathbb{E}(N_b^{(j_0)})\log\mathbb{E}(N_b^{(j_0)})}{\sum_{b\in\mathcal{A}_{j_0},\,\pi_{j_0}([b])=[\widehat{b}]}\mathbb{E}(N_b^{(j_0)})}\\
&= \widetilde{\gamma}_{j_0}\sum_{\widehat{b}\in\mathcal{A}_{j_0+1}}\frac{\nu_{j_0+1}([\widehat{b}])}{\mathbb{E}(N_{\widehat{b}}^{(j_0+1)})}\cdot\sum_{\substack{b\in\mathcal{A}_{j_0}\\ \pi_{j_0}([b])=[\widehat{b}]}}\mathbb{E}(N_b^{(j_0)})\log\mathbb{E}(N_b^{(j_0)})\\
&= \widetilde{\gamma}_{j_0}\sum_{b\in A_{j_0}}\nu_{j_0}([b])\log\mathbb{E}(N_b^{(j_0)})
\end{aligned}$$

by definition of ν_{j_0}. Consequently, noting that $D_{j_0+1}(\nu_{j_0+1}) = D_{j_0}(\nu_{j_0+1},1)$, we get

$$D_{j_0+1}(\nu_{j_0+1}) \le D_{j_0}(\nu_{j_0+1},\widetilde{\theta}_{j_0}) \le D_{j_0}(\widetilde{\nu}_{j_0+1},\widetilde{\theta}_{j_0}) = D_{j_0}. \tag{4.12}$$

Consequently, $D_{j_0+1}\le D_{j_0}$.

Also, the previous inequalities are equalities if and only if the derivative of the convex analytic function $D_{j_0}(\nu_{j_0+1},\cdot)$ vanishes over $[\widetilde{\theta}_k,1]$, hence vanishes everywhere, and $\nu_{j_0+1}=\widetilde{\nu}_{j_0+1}$; notice also that in this case $\sum_{b\in\mathcal{A}_{j_0}}\nu_{j_0}([b])\log\mathbb{E}(N_b^{(j_0)})=0$. This implies that $\nu_{j_0}\in\widetilde{B}_{j_0}$ and $D(\nu_{j_0})=D(\widetilde{\nu}_{j_0})$, so $\widetilde{\nu}_{j_0}=\nu_{j_0}$ hence the Mandelbrot measure $\mu_{\nu_{j_0+1}}$ associated with ν_{j_0+1} coincides with the measure $\mu_{\widetilde{\nu}_{j_0}}$ associated with $\widetilde{\nu}_{j_0}$.

Now, let $j_0+1<j\le k$ and $\nu_j\in\widetilde{\mathcal{B}}_j$. For all $j_0\le i<j$, denote by ν_i the measure $\mathbb{E}(\Pi_{i*}\mu_{\nu_j})$. Remarks 4.4(1) and (2) yield $D_j(\nu_j)\le D_{j_0+1}(\nu_{j_0+1})$ as well as $\sum_{b\in\mathcal{A}_{j_0}}\nu_{j_0}([b])\log\mathbb{E}(N_b^{(j_0)})\le 0$, so again $D_{j_0+1}(\nu_{j_0+1})\le D_{j_0}$. Also, if $j=k$ and $\nu_k\in\widetilde{\mathcal{B}}_{k+1}$, then $D_{k+1}(\nu_k)\le D_k(\nu_k)\le D_{j_0+1}(\nu_{j_0+1})$, since $\sum_{b\in\mathcal{A}_k}\nu_k([b])\log\mathbb{E}(N_b^{(k)})\le 0$. In all these cases, again the same argument as above implies that if there is equality $D_j(\nu_j)=D_{j_0+1}(\nu_{j_0+1})=D_{j_0}$, then the Mandelbrot μ_{ν_j} associated with ν_j coincides with the measure $\mu_{\widetilde{\nu}_{j_0}}$ associated with $\widetilde{\nu}_{j_0}$. In particular, $M^{\vec{\gamma}}$ is uniquely reached. Moreover, we have $i_0=j_0$, and $\theta_{j_0}=\widetilde{\theta}_{j_0}$.

Second case: $\sum_{b\in\mathcal{A}_{j_0}}\widetilde{\nu}_{j_0}([b])\log\mathbb{E}(N_b^{(j_0)})<0$. Since by definition of j_0 the set $\widetilde{\mathcal{B}}_{j_0}$ is not empty, proceeding as when $j_0=k$ we can show that the value D_{j_0} is reached at a measure ν_{j_0} such that $\sum_{b\in\mathcal{A}_{j_0}}\nu_{j_0}([b])\log\mathbb{E}(N_b^{(j_0)})=0$ (notice that again due to Remark 4.4 one has automatically $\sum_{b\in\mathcal{A}_{j_0-1}}\nu_{j_0-1}([b])\log\mathbb{E}(N_b^{(j_0-1)})\le 0$). Thus, we seek for the optimum of $D_{j_0}(\nu_{j_0})$ (see (4.10)) under the constraint

$$\sum_{\widehat{b}\in\mathcal{A}_{j_0+1}}\nu_{j_0+1}([\widehat{b}])\sum_{b\in\mathcal{A}_{j_0}}p(b|\widehat{b})\log\mathbb{E}(N_b^{(j_0)})=0. \tag{4.13}$$

Using the Lagrange multipliers method shows that this approach yields the maximum if and only if there exists a real number θ_0, not depending on ν_{j_0+1}, such that given ν_{j_0+1} one has $p(b|\widehat{b})=p_{\theta_0}(b|\widehat{b})$ (recall the definition (4.11) of $p_\theta(b|\widehat{b})$).

Suppose that such a θ_0 exists. It is then straightforward to check that if the optimum is given by this method it must be reached at $\nu_{\theta_0\phi_{j_0}}$ and equal to $P^{(\widetilde{\gamma}_{j_0},\gamma_{j_0+1},\dots,\gamma_k)}(\theta_0\phi_{j_0},T_{j_0})=$

$P_{j_0}(\theta_0)$. Recall that by (2.7) we have

$$P'_{j_0}(\theta) = \widetilde{\gamma}_{j_0} \sum_{b \in \mathcal{A}_{j_0}} \nu_{\theta_0 \phi_{j_0}}([b]) \log \mathbb{E}(N_b^{(j_0)}).$$

This implies that $\theta_0 \geq \widetilde{\theta}_{j_0}$, since P_{j_0} is convex, (4.13) is equivalent to $P'_{j_0}(\theta_0) = 0$, and our assumption on $\widetilde{\nu}_{j_0}$ asserts that $P'_{j_0}(\widetilde{\theta}_0) < 0$. Also, θ_0 is the infimum of those $\theta \geq \widetilde{\theta}_0$ such that $\sum_{b \in \mathcal{A}_{j_0}} \nu_{\theta \phi_{j_0}}([b]) \log \mathbb{E}(N_b^{(j_0)}) \geq 0$.

Now suppose that $\theta_0 \leq 1$. Then (4.12) holds for any $\nu_{j_0+1} \in \widetilde{B}_{j_0+1}$, with θ_0 instead of $\widetilde{\theta}_{j_0}$ and $\nu_{\theta_0 \phi_{j_0}}$ instead of $\widetilde{\nu}_{j_0}$. Consequently, by the same argument as in the first case we can conclude that D_{j_0} is reached uniquely at $\nu_{\theta_0 \phi_{j_0}}$ and it is not smaller than D_j for $j_0 < j \leq k+1$. Also, $M^{\vec{\gamma}}$ is uniquely reached, $i_0 = j_0$, and $\theta_{i_0} = \theta_0$.

If $\theta_0 > 1$, we see from the definition of the convex function $D_{j_0}(\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}, \cdot)$ and the property $P'_{j_0}(\theta_0) = 0$ that $D'_{j_0}(\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}, \theta_0) = 0$. Thus, by convexity, $D_{j_0+1}(\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}) = D_{j_0}(\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}, 1) \geq D_{j_0} = D_{j_0}(\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}, \theta_0)$, with either strict inequality or the function $D_{j_0}(\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}, \cdot)$ is constant. In the former case, using the definitions we see that $D_{j_0}(\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}, 1) > D_{j_0}$ precisely means

$$\sum_{\widehat{b} \in \mathcal{A}_{j_0+1}} \pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}([\widehat{b}]) \log \mathbb{E}(N_{\widehat{b}}^{(j_0+1)}) > 0.$$

This implies that $\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}} \in \widetilde{\mathcal{B}}_{j_0+1}$ and $D_{j_0} < D_{j_0+1}$. In the latter case, the fact that $D_{j_0}(\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}, \cdot)$ is constant implies, using the expression of this function, that for all $\widehat{b} \in \mathcal{A}_{j_0+1}$ such that $\pi_{j_0 *} \nu_{\theta_0 \phi_{j_0}}([\widehat{b}]) > 0$ the function $\theta \mapsto \sum_{\substack{b \in \mathcal{A}_{j_0} \\ \pi_{j_0}([b]) = [\widehat{b}]}} \mathbb{E}(N_b^{(j_0)})^\theta$ is constant. We conclude that in all these expression $\mathbb{E}(N_b^{(j_0)}) = 1$, which contradicts the fact that $P'_j(\widetilde{\theta}_{j_0}) < 0$. Hence the latter case is empty.

If there is no $\theta \in \mathbb{R}$ such that $P'_{j_0}(\theta) = 0$, then $\mathbb{E}(N_b^{(j_0)}) \leq 1$ for all $b \in \mathcal{A}_{j_0}$, with $\mathbb{E}(N_b^{(j_0)}) < 1$ for some $b \in \mathcal{A}_{j_0}$, and $\mathbb{E}(N_b^{(j_0)}) = 1$ for some other $b \in \mathcal{A}_{j_0}$ (by definition of j_0). Moreover, since a measure ν_{j_0} at which D_{j_0} is reached belongs to $\widetilde{\mathcal{B}}_{j_0}$, we have $\nu_{j_0}([b]) > 0$ only if $\mathbb{E}(N_b^{(j_0)}) = 1$, otherwise we would have $\sum_{b \in \mathcal{A}_{j_0}} \nu_{j_0}([b]) \log \mathbb{E}(N_b^{(j_0)}) < 0$. In particular, this implies that $\nu_{j_0+1}([\widehat{b}]) > 0$ only if $\mathbb{E}(N_{\widehat{b}}^{(j_0+1)}) \geq 1$. optimising $D_{j_0}(\nu_{j_0})$ given ν_{j_0+1} then yields (remind (4.10)), after defining $\widetilde{N}_{\widehat{b}}^{(j_0+1)} = \#\{b \in \mathcal{A}_{j_0} : \Pi_{j_0} b = \widehat{b},\ \mathbb{E}(N_b^{(j_0)}) = 1\}$:

$$p(b|\widehat{b}) = \begin{cases} 1/\widetilde{N}_{\widehat{b}}^{(j_0+1)} & \text{if } \mathbb{E}(N_b^{(j_0)}) = 1, \\ 0 & \text{otherwise} \end{cases}.$$

Consequently,

$$\begin{aligned}
D_{j_0}(\nu_{j_0}) &= h_{\nu_{j_0+1}}^{(\widetilde{\gamma}_{j_0+1},\dots,\gamma_k)}(T_{j_0+1}) + \widetilde{\gamma}_{j_0} \sum_{\widehat{b}\in\mathcal{A}_{j_0+1}:\, \widetilde{N}_{\widehat{b}}^{(j_0+1)}\neq 0} \nu_{j_0+1}([\widehat{b}]) \log \widetilde{N}_{\widehat{b}}^{(j_0+1)} \\
&\leq h_{\nu_{j_0+1}}^{(\widetilde{\gamma}_{j_0+1},\dots,\gamma_k)}(T_{j_0+1}) + \widetilde{\gamma}_{j_0} \sum_{\widehat{b}\in\mathcal{A}_{j_0+1}:\, \widetilde{N}_{\widehat{b}}^{(j_0+1)}\neq 0} \nu_{j_0+1}([\widehat{b}]) \log \mathbb{E}(N_{\widehat{b}}^{(j_0+1)}) \\
&= D_{j_0+1}(\nu_{j_0+1}),
\end{aligned}$$

where the last equality comes from the fact that $\nu_{j_0+1}([\widehat{b}]) > 0$ only if $\mathbb{E}(N_{\widehat{b}}^{(j_0+1)}) \geq 1$. Moreover, this inequality is an equality only if $\mathbb{E}(N_{\widehat{b}}^{(j_0+1)}) = \widetilde{N}_{\widehat{b}}^{(j_0+1)}$ when $\widetilde{N}_{\widehat{b}}^{(j_0+1)} \neq 0$ and $\nu_{j_0+1}([\widehat{b}]) > 0$. We automatically have $\sum_{\widehat{b}\in\mathcal{A}_{j_0+1}} \nu_{j_0+1}([\widehat{b}]) \log \mathbb{E}(N_{\widehat{b}}^{(j_0+1)}) \geq 0$, hence $\nu_{j_0+1} \in \widetilde{\mathcal{B}}_{j_0+1}$. So $D_{j_0} \leq D_{j_0+1}$.

We can now conclude. If $i_0 \leq k$, either $i_0 = j_0$, or $D_{j_0} \leq \ldots \leq D_{i_0-1} \leq D_{i_0}$, and D_{i_0} is reached at $\mu_{\nu_{\widetilde{\theta}_{i_0}\phi_{i_0}}}$. Moreover, the discussions of the first case and the second case when θ_0 exists and belongs to $[\widetilde{\theta}_{j_0}, 1]$ are valid for i_0 and θ_{i_0}, so $M^{\vec{\gamma}}$ is uniquely reached at $\mu_{\nu_{\widetilde{\theta}_{i_0}\phi_{i_0}}}$.

If $i_0 = k+1$, we have $D_{j_0} \leq \ldots \leq D_k$, and we are back to the second case $j_0 = k$, θ_0 exists and $\theta_0 > 1$, or θ_0 does not exists. This yields the desired conclusion. □

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. Also, the assumptions of Theorem 2.2 are fulfilled for μ, and the Hausdorff dimension of μ provides the desired lower bound for $\dim K$.

4.4. **Upper bound for the Hausdorff dimension of K.** Let us start by discussing a first possible attempt to show that $\dim K \leq D_{i_0}$ (recall the definition (4.6)). We could expect to use the measure $\mu = \mu_{\nu_{\theta_{i_0}\phi_{i_0}}}$ of maximal Hausdorff dimension D_{i_0} and show that $\underline{\dim}_{\mathrm{loc}}(\mu, x) \leq D_{i_0}$ everywhere on K; this is the approach used by McMullen as well as Kenyon and Peres in the deterministic case; it 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 $\lim_{p\to\infty} \dim K^{(p)} = D_2$. 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 $\underline{\dim}_{\mathrm{loc}}(\mu, x) \leq D_{i_0}$ everywhere on K; moreover, the sufficient condition just stated to get the sharp upper bound for $\dim K$ 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 $\dim K$. 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 $\dim K$, expressed through a variational principle.

Theorem 4.8. *With probability 1, conditional on $K \neq \emptyset$,*

$$\dim K \leq \inf \Big\{ P_i(\theta) : \, i \in I, \, \widetilde{\theta}_i \leq \theta \leq 1 \Big\}.$$

The sharp upper bound for $\dim K$ follows since by Theorem 4.8, if $i_0 \leq k$, taking $\theta = \theta_{i_0}$ yields the upper bound $\dim K \leq P_{i_0}(\theta_{i_0}) = D_{i_0}$, and if $i_0 = k+1$, $D_{k+1} = P_k(1)$ is an upper bound for $\dim K$ as well.

Before proving Theorem 4.8, 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 $\nu_i \in \mathcal{B}_i$ set

$$D_{i,\theta}(\nu_i) = \widetilde{\gamma}_i \theta \sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \mathbb{E}(N_b^{(i)}) + \widetilde{\gamma}_i h_{\nu_i}(T_i) + \sum_{j=i+1}^{k} \gamma_j h_{\Pi_{i,j\,*}\nu_i}(T_j) \tag{4.14}$$

(in particular $D_{i,\widetilde{\theta}_i}(\nu_i) = D_i(\nu_i)$, recall (2.4) and (4.4)), and for $\rho = (\widetilde{\rho}_i, \rho_i, \ldots, \rho_k) \in \mathcal{B}_i \times \prod_{j=i}^{k} \mathcal{B}_j$, set

$$\widetilde{D}_{i,\theta}(\rho) = \widetilde{\gamma}_i \theta \sum_{b \in \mathcal{A}_i} \widetilde{\rho}_i([b]) \log \mathbb{E}(N_b^{(i)}) + \widetilde{\gamma}_i h_{\rho_i}(T_i) + \sum_{j=i+1}^{k} \gamma_j h_{\rho_j}(T_j). \tag{4.15}$$

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

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

which makes it a compact set. Let

$$\mathcal{R}_i = \{ \rho \in \mathcal{B}_i \times \prod_{j=i}^{k} \mathcal{B}_j : \widetilde{D}_{i,\theta}(\rho) \leq P_i(\theta) \}. \tag{4.16}$$

Ri 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 $x \in X_1$, $2 \le j \le k$, 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 $(n^{-1}\#\{1 \le m \le n : \Pi_j(x)_m = b)_{b\in\mathcal{A}_j}$. Also, let

$$\rho(x,n) = (\widetilde{\rho}_i, \rho_i, \dots, \rho_k), \tag{4.17}$$

where

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

Now, for any $n \in \mathbb{N}^*$ and $U = (U_i, \dots, U_k) \in \mathcal{A}_i^{\ell_i(n)} \times \prod_{j=i+1}^k \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$, we can define $\rho(U) = (\widetilde{\rho}_i(U), \rho_i(U), \dots, \rho_k(U))$ 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 $i+1 \le j \le k$. Note that $\rho_i(U)$ depends on U_i only, so we also denote it by $\rho_i(U_i)$.

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

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

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

$$\begin{aligned}
\rho_i^{(m)}([U_i]) = \prod_{b\in\mathcal{A}_i} \rho_i^{(m)}([b])^{\ell_i(n)\rho_i(U)([b])} &\ge \prod_{b\in\mathcal{A}_i} \rho_i^{(m)}([b])^{\ell_i(n)\rho_i^{(m)}([b])} \prod_{b\in\mathcal{A}_i} \rho_i^{(m)}([b])^{\ell_i(n)\epsilon}\\
&\ge \exp\Big(-\ell_i(n)(h_{\rho_i^{(m)}}(T_i) + \epsilon\log(2/\epsilon))\Big),
\end{aligned}$$

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

Similarly, for each $i+1 \le j \le k$, 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 $|\rho_j(U_j)([b]) - \rho_j^{(m)}([b])| \le \epsilon$ for all $b \in \mathcal{A}_j$ is bounded by $\exp\Big((\ell_j(n) - \ell_{j-1}(n))(h_{\rho_j^{(m)}}(T_j) + \epsilon\log(2/\epsilon))\Big)$. 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

$$\begin{aligned}
\#\mathcal{R}_i(\epsilon,m,n) &\le \prod_{j=i}^m (\#\mathcal{U}_{j,\epsilon,m,n})\\
&\le \exp\Big(\ell_k(n)\epsilon\log(2/\epsilon)\Big)\exp\Big(\ell_i(n)h_{\rho_i^{(m)}}(T_i) + \sum_{j=i+1}^k (\ell_j(n) - \ell_{j-1}(n))h_{\rho_j^{(m)}}(T_j)\Big).
\end{aligned} \tag{4.18}$$

We also notice that if we endow X_i with the metric

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

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 $B_U = \{y \in X_i : y_1 \cdots y_{\ell_i(n)} = U_i, y_{\ell_{j-1}(n)+1} \cdots y_{\ell_j(n)} = U_j \ \forall\, i+1 \le j \le k\}$ 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 $N^{(i)}_{U_{i|\ell_{i-1}(n)}}$ 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 $\Pi_j([V_{\ell_{j-1}(n)+1} \cdots V_{\ell_j(n)}])$, $1 \le j \le i-1$, and B_U determine a unique ball B of X_1 such that $\Pi_i(B) = B_U$. This implies $n_U \le N^{(i)}_{U_{i|\ell_{i-1}(n)}}$. Consequently, for every integer ℓ between $\ell_{i-1}(n)$ and $\ell_i(n)$, we also have $n_U \le N^{(i)}_{U_{i|\ell}}$. In particular,

$$n_U \le N^{(i)}_{U_{i|\lceil \theta \ell_i(n) \rceil}}. \tag{4.19}$$

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 (4.16) of $\mathcal{R}_i$.

Lemma 4.9. *For all $x \in \widetilde{X}_1$:*

(1) $\liminf_{n\to\infty} \widetilde{D}_{i,\theta}(\rho(x,n)) - D_{i,\theta}(\rho_i(x,n)) \le 0$;
(2) $\liminf_{n\to\infty} \widetilde{D}_{i,\theta}(\rho(x,n)) \le P_i(\theta)$;
(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 4.8. It follows from Lemma 4.9(3) that given $\epsilon > 0$, for all $x \in \widetilde{X}_1$, there exists $1 \le m \le M(\epsilon)$ 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:

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

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

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

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

$$\begin{aligned}\mathbb{E}\Big(\mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K)\Big) &\le \sum_{n\ge N} e^{-\frac{n}{\gamma_1}s} \sum_{m=1}^{M(\epsilon)} \sum_{U\in\mathcal{R}_i(\epsilon,m,n)} \mathbb{E}(N^{(i)}_{U_{i|\lceil\theta\ell_i(n)\rceil}}) \\ &= \sum_{n\ge N} e^{-\frac{n}{\gamma_1}s} \sum_{m=1}^{M(\epsilon)} \sum_{U\in\mathcal{R}_i(\epsilon,m,n)} \prod_{\ell=1}^{\lceil\theta\ell_i(n)\rceil} \mathbb{E}(N^{(i)}_{(U_i)_\ell})\end{aligned}$$

and using the definition of $\widetilde{\rho}_i(U)$ to reexpress the right hand side of the last inequality, this yields

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

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

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

We then obtain:

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

Using (4.18), 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 $1\leq m\leq M(\epsilon)$:

$$\begin{aligned}
&(\#\mathcal{R}_i(\epsilon,m,n)) \exp\Big(\lceil\theta\ell_i(n)\rceil \sum_{b\in\mathcal{A}_i} \widetilde{\rho}_i^{(m)}([b])\log\mathbb{E}(N_b^{(i)})\Big)\\
&\leq \widehat{C}_i \exp\Big(\ell_k(n)\epsilon\log(2/\epsilon)\Big)\exp\Big(\frac{n}{\gamma_1}\widetilde{D}_{i,\theta}(\rho^{(m)})\Big)\\
&\leq \widehat{C}_i \exp\Big(\ell_k(n)\epsilon\log(2/\epsilon)\Big)\exp\Big(\frac{n}{\gamma_1}P_i(\theta)\Big) \quad \text{(recall that } \rho^{(m)}\in\mathcal{R}_i).
\end{aligned}$$

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

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

If $s>P_i(\theta)+C_i\epsilon\log(1/\epsilon)$, this yields $\mathbb{E}\Big(\sum_{N\geq 1}\mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K)\Big)<\infty$, so $\lim_{N\to\infty}\mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(K)=0$ and $\dim K\leq s$ almost surely. Since this holds for any fixed $\epsilon>0$ small enough, we get $\dim K\leq P_i(\theta)$ 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 4.8. □

Proof of Lemma 4.9. That (1) ⇒ (2) follows from the fact that $P_i(\theta)=\max\{D_{i,\theta}(\nu_i):\nu_i\in\mathcal{B}_i\}$, and (2) ⇒ (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 $n\in\mathbb{N}^*$,

$$\widetilde{D}_{i,\theta}(\rho(x,n))-D_{i,\theta}(\rho_i(x,n))\leq\epsilon_n+\sum_{j=1}^{J}u_j(\lceil\beta_j n\rceil)-u_j(\lceil\alpha_j n\rceil), \tag{4.20}$$

with $\lim_{n\to\infty}\varepsilon_n = 0$. The desired conclusion is then a direct application of [23, Lemma 5.4], which is a slight extension of the combinatorial lemma used by Kenyon and Peres [34, Lemma 4.1].

To prove (4.20), noting that $\Pi_{i,j_*}\rho_i(x,n) = \rho_j(x,n)$ for all $i \le j \le k$, 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 $w_j(n) = \gamma_j h_{\rho_j(x,n)}(T_j)$:

$$\widetilde{D}_{i,\theta}(\rho(x,n)) - D_{i,\theta}(\rho_i(x,n)) = v_i(\lceil\theta\ell_i(n)\rceil) - v_i(n) + \sum_{j=i}^{k} w_j(\ell_j(n)) - w_j(n)$$
$$+ \sum_{j=i+1}^{k}\gamma_j(h_{\rho_j(T_1^{\ell_{j-1}(n)}x,\ell_j(n)-\ell_{j-1}(n))}(T_j) - h_{\rho_j(x,\ell_j(n))}(T_j)).$$

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, it is easy to see using the definitions and the convexity of $x \ge 0 \mapsto x\log x$ that

$$\begin{aligned} &h_{\rho_j(T_1^{\ell_{j-1}(n)}x,\ell_j(n)-\ell_{j-1}(n))}(T_j) - h_{\rho_j(x,\ell_j(n))}(T_j) \\ &\le \frac{\ell_{j-1}(n)}{\ell_j(n)-\ell_{j-1}(n)}(h_{\rho_j(x,\ell_j(n))}(T_j) - h_{\rho_j(x,\ell_{j-1}(n))}(T_j)). \end{aligned}$$

Setting $\alpha_j = \frac{\widetilde{\gamma}_j}{\gamma_1}$, this implies that

$$\begin{aligned} &\gamma_j(h_{\rho_j(T_1^{\ell_{j-1}(n)}x,\ell_j(n)-\ell_{j-1}(n))}(T_j) - h_{\rho_j(x,\ell_j(n))}(T_j)) \\ &\le \frac{\widetilde{\gamma}_{j-1}}{\gamma_j}(w_j(\lceil\alpha_j n\rceil) - w_j(\lceil\alpha_{j-1}n\rceil)) + o(1). \end{aligned}$$

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 (4.20) holds. □

5. The box counting dimension of K. Proofs of Theorem 2.5 and Corollary 2.7

Proof of Theorem 2.5. 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 [24, 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 [13].

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

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

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

$$E_U = \left\{\exists\,(u_i)_{1\le i\le k} \in \prod_{i=1}^{k}\mathcal{A}_1^{\ell_i(n)-\ell_{i-1}(n)} :\ \text{both } F_k^U(u_1,\dots,u_k) \text{ and } K^{u_1u_2\cdots u_k} \ne \emptyset \text{ hold}\right\}$$

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

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

(note that necessarily $u_1 = U_1$). For $2 \le i \le k$ and $(U_1,\dots,U_i) \in \prod_{j=1}^i \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$, we set

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

(note that $E_1(U_1)$ is simply the event $\{[U_1] \cap K_n \neq \emptyset\}$). We deduce from [24, 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 $2 \le i \le k$ set

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

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

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

Given $(U_1,\dots,U_{i-1}) \in \prod_{j=1}^{i-1} \mathcal{A}_j^{\ell_j(n)-\ell_{j-1}(n)}$, and fixed associated $(u_1,\dots,u_{i-1})$ such that $F_{i-1}^{(U_1,U_2,\dots,U_{i-1})}(u_1,\cdots,u_{i-1})$ holds, following the arguments of [24], 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 $u_i \in \mathcal{A}_1$ such that $F_i^{(U_1,U_2,\dots,U_i)}(u_1,\cdots,u_i)$ holds, is a random variable $Z_{i,\ell_i(n)-\ell_{i-1}(n)}(U_1,\dots,U_{i-1})$ so that the $Z_{i,\ell_i(n)-\ell_{i-1}(n)}(U_1,\dots,U_{i-1})$ are independent, and identically distributed. Moreover, setting $\widetilde{\ell}_i(n) = \ell_i(n) - \ell_{i-1}(n)$, one has both $\lim_{n\to\infty} \frac{\log \mathbb{E}(Z_{i,\tilde{\ell}_i(n)})(U_1,\dots,U_{i-1})}{n} = \frac{\gamma_i}{\gamma_1}\psi_i(\widehat{\theta}_i) > 0$ and, conditional on $K^{u_1\cdots u_{i-1}} \neq \emptyset$, $\lim_{n\to\infty} \frac{\log(Z_{i,\tilde{\ell}_i(n)}(U_1,\dots,U_{i-1})}{n} = \frac{\gamma_i}{\gamma_1}\psi_i(\widehat{\theta}_i)$ almost surely. Then, again the same reasoning as in [24] 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 $n \ge 1$ 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 $U = (U_1,\dots,U_k) \in \prod_{i=1}^k \mathcal{A}_i^{\ell_i(n)-\ell_{i-1}(n)}$ and associated $(u_1 = U_1, u_2, \dots, u_k) \in$

$\prod_{i=1}^k \mathcal{A}_1^{\ell_i(n)-\ell_{i-1}(n)}$ such that $F_k^{(U_1,\dots,U_k)}(u_1,\dots,u_k)$ hold. The events $\{K^{u_1u_2\cdots u_k}\neq\emptyset\}$ being independent, with the same probability $\mathbb{P}(K\neq\emptyset)$, and independent of the events $F_k^{(U_1,U_2,\dots,U_k)}(u_1,u_2,\dots,u_k)$, using [24, 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. $\dim_B K$. □

Next we state, using our notations, 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 $2\le i\le k$. With probability 1, conditional on $K\neq\emptyset$,*

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

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 2.7. It results from (2.3) and Proposition 5.1 that for $\dim K=\dim_B(K)$ 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 $2\le i\le k$, 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 $\dim_e^{\vec{\gamma}}(\mu)=\dim_B K$. □

6. Projections of K and μ to factors of X_1. Proofs of Theorems 2.8, 2.9 and 2.11, and Corollary 2.12

The proofs will be sketched.

Sketch of the proof of Theorem 2.8. 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. □

Sketch of the proof of Theorem 2.9. We start with the lower bound. We redefine $\widetilde{\mathcal{B}}_i$ as $\widetilde{\mathcal{B}}_i = \Big\{\nu_i \in \mathcal{B}_i : \sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \mathbb{E}(N_b^{(i)}) \geq 0\Big\}$, and we set

$$\overline{D}_i = \max\{h_{\nu_i}^{\vec{\gamma}^i}(T_i) : \nu_i \in \widetilde{\mathcal{B}}_i\}.$$

Here, by convention, $\max(\emptyset) = -\infty$.

Arguing like in Section 4.2, we can get that a lower bound for $\dim \Pi_i(K)$ is given by $\max(\overline{D}_i, D_{i+1}, \ldots, D_{k+1})$. We set $k_0 = \min\{j \geq i : \exists\, b \in \mathcal{A}_j, \mathbb{E}(N_b^{(j)}) \geq 1\}$, with $\min(\emptyset) = k+1$. Clearly $\max(\overline{D}_i, D_{i+1}, \ldots, D_{k+1}) = \max(D_{k_0}, \ldots, D_{k+1})$ if $k_0 > i$.

Suppose that $k_0 = i$. We know that $h_{\nu_i}^{\vec{\gamma}^i}(T_i)$ reaches its maximum $P_i(0)$ at $\nu_{0\cdot\phi_i}$, and $\nu_{0\cdot\phi_i} \in \widetilde{B}_i$ if and only if $P_i'(0) = \sum_{b \in \mathcal{A}_i} \nu_{0\cdot\phi_i}([b]) \log \mathbb{E}(N_b^{(i)}) \geq 0$.

Suppose, moreover, that $P_i'(0) \geq 0$. In this case, it is easily seen that $\overline{D}_i$ is not smaller than D_j, $i+1 \leq j \leq k+1$. If $P_i'(0) > 0$ then it is also easy to construct infinitely many Mandelbrot measures μ which share with $\mu_{\nu_{0\cdot\phi_i}}$ the property that $\mathbb{E}(\Pi_{i*}\mu) = \nu_{0\cdot\phi_i}$ and $\dim_e(\mu) > h_{\nu_{0\cdot\phi_i}}$. On the contrary, if $P_i'(0) = 0$, $\mu_{\nu_{0\cdot\phi_i}}$ is the unique Mandelbrot measure μ such that $\dim_e^{\vec{\gamma}^i}(\mu) = P_i(0)$, and we let the reader check that no other Mandelbrot measure of the form μ_{ν_j} with $j > i$ and $\nu_j \in \widetilde{B}_j$ is such that $\dim_e^{\vec{\gamma}_i}(\mu_{\nu_j}) = P_i(0)$. Note that in any case $j_0 = i$ and $\theta_{j_0} = 0$.

Suppose now that $k_0 = i$ and $P_i'(0) < 0$. To get $\overline{D}_i$ we must maximize $h_{\nu_i}^{\gamma_i}(T_i)$ over those $\nu_i \in \mathcal{B}_i$ such that $\sum_{b \in \mathcal{A}_i} \nu_i([b]) \log \mathbb{E}(N_b^{(i)}) = 0$. Here we meet a situation similar to that we discussed in the proof of Lemma 4.6. The only difference is that $\widetilde{\theta}_{j_0}$ is replaced here by 0. It turns out that either there exists $\theta \in [0,1]$ such that $P_i'(\theta) = 0$ and $\max(\overline{D}_i, D_{i+1}, \ldots, D_{k+1}) = P_i(\theta)$, or $\max(\overline{D}_i, D_{i+1}, \ldots, D_{k+1}) = P_i(\theta) = \max(D_{i+1}, \ldots, D_{k+1})$. In the former case, we also have that $\mu_{\nu_{\theta\phi_i}}$ is the unique Mandelbrot measure μ such that $\dim_e^{\vec{\gamma}_i}(\mu) = P_i(\theta)$, $j_0 = i$ and $\theta_{j_0} = \theta$. In the latter case, we are back to the discussion of the proof of Lemma 4.6 and we also get the desired conclusion.

If $k_0 > i$, since we seek for $\max(D_{k_0}, \dots, D_{k+1})$, the situation also boils down to that of Lemma 4.6.

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

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

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 (4.14) and (4.15)) to $\theta \in [0,1]$ and for $j = i$ we redefine the vector $\rho(x,n)$ of (4.17) by taking $\widetilde{\rho}_i = \rho_i(x, \ell_i(n))$. It is readily seen from the proof of Lemma 4.9 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 4.8, for each $j \in \{i, \dots, k\}$, 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 $\mathbb{E}\big((N_{U_i}^{(i)})^\theta\big) \leq \mathbb{E}\big(N_{U_i}^{(i)}\big)^\theta$, we can use similar estimates as in the proof of Theorem 4.8 to now estimate $\mathcal{H}^s_{e^{-\frac{N}{\gamma_1}}}(\Pi_i(K)))$, and this yields $\dim \Pi_i(K) \leq P_i(\theta)$ (here we followed the same idea as that used in [17] to deal with projections of planar statistically self-similar limit sets of fractal percolation).

Next, suppose $j \in \{i+1, \dots, k\}$ 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 4.8 yield $\dim \Pi_i(K) \leq P_i(\theta)$ for all $\theta \in [\widetilde{\theta}_j, 1]$. □

Sketch of the proof of Theorem 2.11. This is similar to the proof of Theorem 2.5, 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

$$E^i(U) = \Big\{\exists\, (u_j)_{i \leq j \leq k} \in \mathcal{A}_1^{\ell_i(n)} \times \prod_{j=i+1}^{k} \mathcal{A}_1^{\ell_j(n) - \ell_{j-1}(n)} : \\ \text{both } F_k^{i,(U_i,\dots,U_k)}(u_i, \dots, u_k) \text{ and } K^{u_i u_2 \cdots u_k} \neq \emptyset \text{ hold}\},$$

and where

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

One deduces easily from [14] (see alternatively [17] 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 2.5 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 2.12. If $\widehat{\theta}_i=1$, using Proposition 5.1 we see that the equality between $\dim\Pi_i(K)$ and $\dim_B\Pi_i(K)$ imposes that $\dim\Pi_i(K)$ is attained by the branching measure, and the situation boils down to that of Corollary 2.7. This gives point (1) of the statement.

Suppose now that $\dim\Pi_i(K)=\dim_B\Pi_i(K)$, $\widehat{\theta}_i<1$ and $\psi_i'(\widehat{\theta}_i)=0$ (which is automatically true if $0<\widehat{\theta}_i<1$). The equality between $\dim\Pi_i(K)$ and $\dim_B\Pi_i(K)$ imposes that if μ stands for the unique Mandelbrot measure supported on K such that $\dim_e^{\vec{\gamma}_i}(\mu)=\dim\Pi_i(K)$, 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 $j\in I_i(=\{i,\dots,k\})$ such that $j\le j_0'$, we must have $\Pi_{i,j*}\nu_i=\nu_{j,\widehat{\theta}_j}$. Using that for all $b\in\widetilde{A}_j$, 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

$$\begin{aligned}0=\psi_i'(\widehat{\theta}_i)-\psi_j'(\widehat{\theta}_j)&=\sum_{b'\in\widetilde{\mathcal{A}}_i}\nu_{i,\widehat{\theta}_i}([b'])\log\mathbb{E}(N_{b'}^{(i)})-\sum_{b\in\widetilde{\mathcal{A}}_j}\nu_{j,\widehat{\theta}_j}([b])\log\mathbb{E}(N_b^{(j)})\\&=\sum_{b\in\widetilde{\mathcal{A}}_j}\sum_{b'\in\Pi_{i,j}^{-1}(b)}\nu_{i,\widehat{\theta}_i}([b'])\log\frac{\mathbb{E}(N_{b'}^{(i)})}{\mathbb{E}(N_b^{(j)})}.\end{aligned}$$

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 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 2.9 and 2.11 that $\dim\Pi_i(K)=\dim_B\Pi_i(K)$.

The last case easily follows from the previous discussion. □

7. Dimension of conditional measures. Proof of Theorem 2.14

Since the proof of Theorem 2.15 is very similar to that of Theorem 2.14, 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 2.8 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 2.13 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 of not.

When $\mu_\omega = \mu \neq 0$, for $\Pi_{i*}\mu_\omega$-almost every z, the conditional measure μ^z_ω is supported on $K^z = \pi^{-1}(\{z\}) \cap K$, obtained as the weak-star limit, as $n \to \infty$, of the measures $\mu^z_{\omega,n}$ 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 $\mu^z_\omega([J]) = \lim_{n\to\infty} \frac{\mu_\omega([J]\cap\Pi_i^{-1}([z_{|n}]))}{\Pi_{i*}\mu_\omega([z_{|n}])}$.

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 $\widetilde{\mu}^z_{\omega,n} = f_{\omega,n}(z)\mu^z_{\omega,n}$ weakly converges to the measure $\widetilde{\mu}^z_\omega$ defined as $f_\omega(z)\mu^z_\omega$.

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 $\mu^x_\omega = \widetilde{\mu}^x_\omega = 0$.

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

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

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 (3.2), we also have

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

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 (2.11) 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

$$\mathbb{E}_{\mathbb{P}\otimes\nu_i}\Big(\sum_{B_U \in \mathcal{F}_n} \widetilde{\mu}^z(B_U)^q\Big) \leq C_q \exp\Big(-\frac{n}{\gamma_1}(q-1)D + O((q-1)^2))\Big). \tag{7.1}$$

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 $\widetilde{\mu}^z_\omega$ has been defined equal to 0. By Fatou's lemma, we have

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

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

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

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:

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

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

$$m(U_1,\dots,U_{i-1},L,L') = \sum_{\substack{(J_1,\dots,J_{i-1}):\, \Pi_{j,i}(J_j)=U_j,\\ J=J_i\cdots J_k:\, \Pi_i(J_j)=L_j}} \sum_{J':\, \Pi_i(J')=L'} \frac{\mu([J_1\cdots J_{i-1}J'])}{\nu_i([L_1\cdots L_{i-1}L'])}. \tag{7.2}$$

Suppose that $q \geq 1$. 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

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

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 $R_{n,p} = \mathbb{E}_{\mathbb{P}\otimes\nu_i}(X^q_{p+\ell_k(n)-\ell_{i-1}(n)})$, 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.

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

(7.3)

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

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

$$S_{j,n} = \mathbb{E}\sum_{U_j} \phi_j(U_j)\Big(\sum_{J_j:\, \Pi_j(J_j)=U_j} \frac{\mu_{\ell_j(n)-\ell_{j-1}(n)}([J_j])}{\nu_j([U_j])}\Big)^q, \tag{7.4}$$

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

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

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

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

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

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}\otimes\nu_i} \sum_{B_U\in\mathcal{F}_n} \widetilde{\mu}^z(B_U)^q &\leq C_q \exp\Big(q-1)(\ell_{j_0-1}(n)(h_{\nu_i}(T_i) - \dim_e(\mu)) \\
&\qquad + (q-1)\sum_{j=j_0}^{i-1} (\ell_j(n) - \ell_{j-1}(n))(h_{\nu_i}(T_i) - h_{\nu_j}(T_j)) + O((q-1)^2)n\Big) \\
&= C_q \exp\Big(-\frac{n}{\gamma_1}(q-1)D + O((q-1)^2))\Big),
\end{aligned}$$

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

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

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

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

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

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

In our main statements about the Hausdorff and box-counting dimension of K and its projections for simplicity we assumed all the γ_i, $2 \le i \le k$ to be positive, which in the Euclidean realisation of Section 2.6 corresponds to $m_1 > \cdots > m_k \ge 2$. It turns out that up to slight modifications in the statement and proofs, our results cover the general configuration $m_1 \ge \cdots \ge m_k \ge 2$, 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 $1 \le i < j \le k$, $x \in X_i$ and $n \ge 1$, for the induced metric by $d_{\tilde{\gamma}}$ 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 4.8, the only change to make to cover the case $\gamma_i \ge 0$ for all $2 \le i \le k$ is to set $I = \{2 \le i \le k : \gamma_i > 0$ and replace k by $\sup(I)$ in (2.9). The proofs adapt readily.

For the statements of Section 2.4, one has to replace I_i by $\{i\} \cup \{i < j \le k : \gamma_j > 0\}$ and replace k by $\sup(I_i)$ in (2.10). Again, the modifications in the proofs are left to the reader.

Acknowledgements. The research of both authors was supported in part by University of Paris 13, the HKRGC GRF grants (projects CUHK401013, CUHK14302415), and the France/Hong Kong joint research scheme PROCORE (33160RE, F-CUHK402/14).

References

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

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

[3] B. Barany, M. Hochman, A. Rapaport, Hausdorff dimension of planar self-affine sets and measures, arXiv:1712.07353.

[4] B. Barany, M. Rams, K. Simon, On the dimension of triangular self-affine sets, Ergod. Th. & Dyn. 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] Bedford, T. 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, J. Th. Nomb. Bordeaux, **4** (1992), 129–140.

[10] F. Ben Nasr, Ensembles alatoires 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] T. Das, D. Simmons, The Hausdorff and dynamical dimensions of self-affine sponges: a dimension gap result, Inv. Math., **210** (2017), 85–134.

[13] M. Dekking, On the survival probability of a branching process in a finite state i.i.d. environment, **27** 1987, 151–157.

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

[15] Durrett, R., Liggett, T. (1983). Fixed points of the smoothing transformation. *Z. Wahrsch. Verw. Gebiete* **64** 275–301.
[16] K. J. Falconer, Dimensions and measures of quasi self-similar sets, *Proc. Amer. Math. Soc.*, **106** (1989, 543–554.
[17] K. J. Falconer, Projections of random Cantor sets, J. Theoret. Probab., **2** (1989), 65–70.
[18] K. J. Falconer, X. Jin, Exact dimensionality and projections of random self-similar measures and sets, J. London Math. Soc., **90** (2014), 388–412.
[19] K. Falconer, T. Kempton, Planar self-affine sets with equal Hausdorff, box and affinity dimensions, Ergod.Th. & Dynam. Sys., **38** (2018), 1369–1388.
[20] K. Falconer and J. Miao. Dimensions of self-affine fractals and multifractals generated by upper-triangular matrices. Fractals, 15(3):289299, 2007.
[21] A.-H. Fan, K.-S. Lau, H. Rao, Relationships between different dimensions of a measure, Monatsh. Math., **135** (2002), no. 3, 191–201.
[22] Dimension of invariant measures for self-affine iterated function systems, arXiv:1901.01691v1.
[23] D. J. Feng, W. Huang, Variational principle for weighted topological pressure, J. Math. Pures Appl., **106** (2016), 411–452
[24] D. Gatzouras, S. Lalley, Statistically self-affine sets: Hausdorff and box dimensions, J. Theoretical Probability, **7** (1994), 437–468.
[25] D. Gatzouras, Y. Peres, Invariant measures of full dimensions for some expanding maps, Ergod. Th. & Dyn. Sys., **17** (1997), 147–167.
[26] S. Graf, R. D. Mauldin, S. C. Williams, The exact Hausdorff dimension in random recursive constructions, 1988. Amer. Math. Soc., Providence, RI.
[27] M. Hochman, Dimension theory of self-similar sets and measures, Proc. Int. Cong. of Math. 2018, Rio de Janeiro, Vol. 2 (1943–1966).
[28] M. Hochman, A. Rapaport, Hausdorff dimension of planar self-affine sets and measures with overlaps, arXiv:1904.09812.
[29] J. Hutchinson, Fractals and self-similarity, Indiana Univ. Math. J., **30** (1981) 713–747.
[30] T. Jordan, M. Pollicott and K. Simon, Hausdorff dimension for randomly perturbed self affine attractors. *Comm. Math. Phys.* **270** (2007), 519-544.
[31] A. Käenmäki, On natural invariant measures on generalised iterated function systems. *Ann. Acad. Sci. Fenn. Math.* **29** (2004), 419–458.
[32] A. Käenmäki, M. Vilppolainen, Dimension and measures on sub-self-affine sets. *Monatsh. Math.*, **61** (2010), no. 3, 271–293.
[33] J.-P. Kahane, Multiplications aléatoires et dimensions de Hausdorff, Ann. Inst. Henri Poincaré Probab. Stat., **23** (1987), 289–296.
[34] J.-P. Kahane and J. Peyrière, Sur certaines martingales de B. Mandelbrot. Adv. Math., **22** (1976), 131–145.
[35] R. Kenyon and Y. Peres, Measures of full dimension on affine-invariant sets, *Ergod. Th.* & *Dynam. Sys.* **16** (1996), 307–323.
[36] Y. Kifer, Fractal dimensions and random transformations, Trans. AMS, **348** (1996), 2003-2038.
[37] I. Kolossvary, K. Simon, Triangular Gatzouras-Lalley-type planar carpets with overlaps, arXiv:1807.09717.
[38] S. Lalley, D. Gatzouras, Hausdorff and box dimensions of certain self-affine fractals, Indiana Univ. Math. J., **41** (1992), 533–568.
[39] 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.
[40] C. McMullen, The Hausdorff dimension of general Sierpinsky carpets, *Nagoya Math. J.* **96** (1984), 1–9.
[41] S.-M. Ngai, A dimension result arising from the L^q-spectrum of a measure, Proc. Amer. Math. Soc., **125** (1997), 2943–2951.
[42] J. Peyrière, Comparaison de deux notions de dimension, Bull. S.M.F., **114** (1986), 97–103.
[43] 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.
[44] B. Solomyak, Measure and dimension for some fractal families. *Math. Proc. Cambridge Philos. Soc.* **124** (1998), 531–546.

[45] 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.
[46] T. Watanabe, The Hausdorff measure on the boundary of a Galton-Watson tree, Ann. Probab., **35** (2007), 1007-1038.

Laboratoire de Géométrie, Analyse et Applications, CNRS, UMR 7539, Université Sorbonne Paris Nord, CNRS, UMR 7539, F-93430, Villetaneuse, France

E-mail address: barral@math.univ-paris13.fr

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

E-mail address: djfeng@math.cuhk.edu.hk