

BI-LIPSCHITZ EMBEDDING OF ULTRAMETRIC CANTOR SETS INTO EUCLIDEAN SPACES

JEAN V. BELLISSARD, ANTOINE JULIEN

ABSTRACT. An ultrametric Cantor set can be seen as the boundary of a rooted weighted tree called the Michon tree. The notion of Assouad dimension is re-interpreted as seen on the Michon tree. The Assouad dimension of an ultrametric Cantor set is finite if and only if the space is bi-Lipschitz embeddable in a finite dimensional Euclidean space. This result, due to Assouad and refined by Luukkainen–Movahedi–Lankarani is re-proved in the Michon tree formalism. It is applied to answer the embedding question for some spaces which can be seen naturally as boundary of trees: linearly repetitive subshifts, Sturmian subshifts, and the boundary of Galton–Watson trees with random weights. Some of these give examples of nonembeddable spaces with finite Hausdorff dimension.

1. INTRODUCTION

This article is dedicated to giving a concrete description of the bi-Lipschitz embedding theorems for ultrametric Cantor sets in view of their applications to various situations like the case of tiling spaces or to the boundary of rooted weighted trees. In particular several important examples will be given of ultrametric Cantor sets with finite Hausdorff dimension that are not bi-Lipschitz embeddable in a finite dimensional Euclidean space.

The problem of embedding metric spaces in the standard Euclidean space \mathbb{R}^n is fairly old, and has many faces. Hence Fréchet’s Theorem [15] shows that any finite metric space with $n + 1$ elements can be embedded isometrically in \mathbb{R}^n when equipped with the ℓ^∞ -norm. Schoenberg [31, 32] gave a necessary and sufficient condition for a finite metric space to be embeddable isometrically in the Euclidean space \mathbb{R}^n . These works are commonly used in data analysis today. Beyond finite metric space, the case of compact ultrametric spaces is probably the simplest to consider next. An ultrametric space is a metric space where the distance satisfies the strengthened triangle inequality: $d(x, z) \leq \max\{d(x, y), d(y, z)\}$ for any three points x, y, z in the space. It is known [29] that any such space is isometrically embeddable in a (real) Hilbert space.

However the requirement that the embedding is isometric is too strong in general. For instance an ultrametric space is unlikely to be isometrically identified with a subset of a finite dimensional Euclidean space. The condition that the embedding should be isometric has to be weakened. A map $f : (X, d) \rightarrow (\mathbb{R}^n, \|\cdot\|)$ is a bi-Lipschitz embedding if it is a homeomorphism onto its image, and furthermore

Work supported in part by NSF grants DMS 0901514, DMS 1160962 and by the CRC701, *Spectral Structures and Topological Methods in Mathematics*, Mathematik, Universität Bielefeld, Germany. Substantial part of this work was made while the second author was employed by the University of Victoria, with the support of the Pacific Institute for the Mathematical Sciences.

$$\exists c \geq 1, \quad \forall x, y \in X, \quad \frac{1}{c} d(x, y) \leq \|f(x) - f(y)\| \leq c d(x, y).$$

In the rest of this article, an *embedding* will always be assumed to be bi-Lipschitz, and a space will be called embeddable (or f -embeddable, f standing for “finite-dimensional”) if it can be embedded into some \mathbb{R}^n equipped with its Euclidean metric.

The concept of Assouad dimension has been a breakthrough in the problem of embeddability [3, 4, 5] (see also Semmes [33]). A space has finite Assouad dimension if and only if it is doubling, that is, for some $M \in \mathbb{N}$, any ball can be covered by at most M balls of half radius. Moreover a separable metric space is f -embeddable only if its Assouad dimension is finite. If, in addition, the space is ultrametric, then its f -embeddability is equivalent to having finite Assouad dimension. This result, obtained by Assouad, was strengthened by Luukkainen and Movahedi-Lankarani [25] and then Luosto [24]: an ultrametric space is bi-Lipschitz embeddable in \mathbb{R}^n if and only if its Assouad dimension is strictly less than n . The Assouad dimension seems therefore perfectly fitted to answer the embedding question for ultrametric spaces.

In this paper, the focus is on ultrametric Cantor sets. Cantor sets are totally disconnected metrizable compact spaces, without isolated points. A theorem of Brouwer [8] states that any two such sets are homeomorphic (hence any of these sets can be called “*the* Cantor set”). However, the Cantor set can carry a lot of different ultrametrics. A classification of all ultrametric on a Cantor set was given by Michon [27]: an ultrametric Cantor set can be represented by a rooted infinite tree, and the ultrametric is described by weights on the vertices (which correspond to the diameters of a basis of neighborhoods).

This paper is structured as follows: Section 2 gives the basic notations and terminology. In particular the definition of Assouad dimension of a metric space and the description of ultrametric Cantor sets arising as boundaries of weighted rooted trees are described (which by Michon’s Theorem, are all ultrametric Cantor sets in existence).

Section 3 is dedicated to existing results about bi-Lipschitz embeddings. The interpretation of the Assouad dimension in terms of the Michon tree is discussed. Luukkainen–Movahedi-Lankarani’s result is re-proved in this formalism and takes a very concrete form. The proof is relatively short, and self-sustained. It also has the advantage that the embedding seems slightly more explicit than in the original proof.

As a corollary (see [19]), this result shows that the classical quasicrystalline tilings, like the Penrose tiling, the octagonal tiling and the three classes of icosahedral tilings used to model quasicrystalline metallic alloys [18] are embeddable when equipped with the combinatorial metric. Actually if d is the dimension of such a tiling, the Hausdorff dimension of its tiling space is equal to d and is embeddable in \mathbb{R}^{d+1} , showing that the Assouad dimension is larger than or equal to d and smaller than $d + 1$. It is likely that, in such cases, the Assouad dimension is actually equal to d .

Section 4 gives a class of examples defined through a special type of subshift. An ultrametric Cantor set appears as the transversal of quasiperiodic tilings (or in one dimension, as subshifts of $\{0, 1\}^{\mathbb{Z}}$). The following two theorems are proved.

Theorem 1. *The tiling space of a (bi-infinite, one-dimensional) linearly repetitive sequence is f -embeddable.*

Section 5 investigates the case of Sturmian sequences leading to

Theorem 2. *The tiling space of a Sturmian sequence x associated with the irrational number $\alpha \in (0, 1)$ is f -embeddable if and only if α has bounded type, that is if and only if its continued fraction expansion has bounded partial quotients. In particular for Lebesgue-almost every α this tiling space is not f -embeddable.*

Interestingly enough the Minkowski (or box-counting) dimension of a subshift is related with its complexity (see [19, Corollary 3.15]). However, its f -embeddability (and hence its Assouad dimension) has more to do with the recurrence properties of the orbits, and therefore reflects some of the dynamical properties of the subshift. Both combinatorial and dynamical properties of the subshift seem to be captured by the metric. This remark is one of the motivations for studying the metric properties of ultrametric Cantor sets arising from tilings: the distance provides an additional structure which reflects dynamical or combinatorial properties of the tiling. A notion of spectral dimension associated with a family of spectral triples defined on this space had been investigated by Pearson and the first author [30]. It can be seen here that each different concepts of dimension give a specific perspective on these spaces.

In Section 6, the Hausdorff dimension is discussed in terms of the Michon tree representation. In a previous paper, Savinien and the second author [19] showed that a *self-similar* ultrametric Cantor set is embeddable in \mathbb{R}^n as soon as n is strictly more than the Hausdorff dimension of the space. However, in general, the Hausdorff dimension has no reason to be a good predictor of the embeddability of a space. The Hausdorff dimension is also discussed to prepare the last example discussed in this paper.

Section 7 concerns random trees obtained from a Galton–Watson branching process [16, 6]. Let p be a probability defined on the set $\mathbb{N} = \{0, 1, 2, \dots\}$ of natural integers. A *GW*-tree is defined as a rooted tree in which each vertex v has M_v children, where the M_v are independent, identically distributed random variables with probability distribution p (p is a distribution on the natural integers). It is a classical result in probability, partly proved initially by the Reverend Watson [34], called the *Galton–Watson–Haldane–Steffensen critical Theorem* [21], that if the average number of children $\langle M_v \rangle$ is less than or equal to one, the tree obtained in this way is almost surely finite (extinction). In particular its reduced tree is empty. On the other hand, as was eventually proved by Steffensen in 1930, if $\langle M_v \rangle > 1$, the probability of extinction is less than one. However, the random tree produced in this way is likely to have dangling vertices. One way to avoid such a property is to force every vertex to have at least two children. This can be done by demanding that $p_0 = p_1 = 0$, namely that p be supported by $[2, \infty) \subset \mathbb{N}$: then the Galton–Watson tree is automatically reduced and this model can be called a *reduced Random tree*. A first result is the following:

Proposition 1. *Let p be a probability on the set $[2, \infty) \subset \mathbb{N}$. If it has an infinite support, the boundary of the reduced Galton–Watson tree associated with this probability is almost surely not f -embeddable.*

In such a case a *random weight* can be added in the following way: a family $(\lambda_v)_{v \in \mathcal{V}}$ of *i.i.d.* random variables supported in $[0, 1]$, with common distribution ρ is defined. A weight is defined inductively by setting that the weight of the root is equal to 1, and $\kappa(v) = \lambda_v \kappa(u)$ if v is a child of u . In order that this defines a weight, it is required that $\rho\{0\} = 0$ and $\rho\{1\} < 1$. Then

Theorem 3. *Let p be a probability on the set $[2, \infty) \subset \mathbb{N}$ and let ρ be a probability on $[0, 1]$ such that $\rho\{0\} = 0$ and $\rho\{1\} < 1$. If $m = \langle M_v \rangle < \rho\{1\}^{-1}$, let $s = s_m$ be the unique solution of $\langle \lambda^s \rangle m = 1$. Then the reduced random tree \mathcal{T} produced by the Galton–Watson process associated with p and endowed with the random weight associated with ρ gives rise to an ultrametric Cantor*

set $\partial\mathcal{T}$ with Hausdorff dimension s_m almost surely. In addition, its Hausdorff measure exists almost surely and is a random probability measure.

Note that the Hausdorff dimension had been computed previously when the random variables λ_v are deterministic (for example equal to e^{-1} or $1/2$), see Hawkes [17].

Several observations can be made as a result of the analysis carried out in this paper. First, the Sturmian and Galton–Watson trees provide very natural examples of spaces which have finite Hausdorff dimension, but infinite Assouad dimension. More specifically, given a weighted tree in which the weights decrease exponentially along the branches, the Hausdorff dimension of its boundary is related to its *average* number of children per vertex, while its Assouad dimension is related to the *maximum* number of children per vertex.

As a final anecdote, real trees growing in nature satisfy the requirement that the length of the branches decreases exponential fast with the generation rank. Hence they behave like embeddable trees (in this case they are embedded in \mathbb{R}^3) if the weight is a measure of the length of their branches. The previous results give strong constraints about the growth of real trees since the Assouad dimension is at most 3.

Acknowledgements: J. B. thanks A. Grigor’yan and his group at the CRC701, University of Bielefeld for giving him the opportunity to present this result prior to publication. A. J. thanks J. Savinien for discussions on this embedding problem.

2. DEFINITIONS AND NOTATIONS

2.1. Trees and ultrametric Cantor sets. In this paper, a rooted tree \mathcal{T} is described by a set of vertices \mathcal{V} , which is countable and satisfies the following properties:

- (i) \mathcal{V} is partitioned into $\mathcal{V} = \bigsqcup_{n \geq 0} \mathcal{V}_n$, and a predecessor (or “parent”) map is defined $\mathcal{V}_{n+1} \rightarrow \mathcal{V}_n$ for all $n \geq 1$,
- (ii) \mathcal{V}_0 consists of a unique element, called the root and noted \bullet .

Given a vertex v , any vertex which admits v as a parent is called a child of v . If $v \in \mathcal{V}_n$, the integer n is noted $|v|$ and called the *depth* or the *generation* of v . There is a relation between vertices: $v \succeq w$ if there is a sequence of vertices $\gamma = (v_0, v_1, \dots, v_n)$, with $v = v_0$ and $w = v_n$, such that each v_k is the parent of v_{k+1} . This binary relation reads: “ v is an ancestor of w ” or “ w is a descendant of v ”. It is an order relation with the root as maximal element. In this case, γ is called a *path* from v to w , which *goes through* vertices v_0, v_1, \dots, v_n . A vertex is called *dangling* whenever it has no children, and *branching* if it has at least two children.

It is sometimes convenient to encode the “parent–child” relation by means of edges (this is how trees are defined in graph theory): a tree can be given as a set of vertex and a set of edges \mathcal{E} , which are both partitioned as above, with two maps $r : \mathcal{E}_n \rightarrow \mathcal{V}_{n+1}$ and $s : \mathcal{E}_n \rightarrow \mathcal{V}_n$ (range and source respectively), such that r is bijective. Then the parent map is simply $s \circ r^{-1}$. These definitions are equivalent (note in particular that a path can be defined by a sequence of edges or a sequence of vertices).

The trees considered here are furthermore supposed to be locally finite (every vertex has finitely many children), and without dangling vertices. Under these assumptions, there is a well-behaved topological space, called the boundary of the tree, and defined as

$$\partial\mathcal{T} = \left\{ (v_0, v_1, \dots) \in \prod_{n \in \mathbb{N}} \mathcal{V}_n ; v_0 = \bullet ; v_k \text{ is the parent of } v_{k+1} \right\} \stackrel{def}{=} \varprojlim (\mathcal{V}_n, \rightarrow),$$

which is nothing but the definition of the inverse limit of the sets \mathcal{V}_n under the parent map $\mathcal{V}_{k+1} \rightarrow \mathcal{V}_k$. Under the local finiteness condition, all \mathcal{V}_n are finite. Therefore, equipped with the product topology, $\partial\mathcal{T}$ is a totally disconnected compact set, as an inverse limit of finite sets. Given a vertex v , let $[v]$ be the set of all element $(v_n)_{n \in \mathbb{N}}$ which pass through v . Then $\{[v] ; v \in \mathcal{V}\}$ is a basis for the topology. A rooted tree is called *Cantorian* whenever, in addition, each vertex admits a branching descendant. The boundary of a Cantorian tree has no isolated points, and is therefore a Cantor set.

A *weight* on the rooted tree \mathcal{T} is a map $\kappa : \mathcal{V} \rightarrow (0, \infty)$ such that

- (i) if $w \preceq v$ then $\kappa(w) \leq \kappa(v)$,
- (ii) $\lim_{|v| \rightarrow \infty} \kappa(v) = 0$.

A weight κ induces an ultrametric d_κ on $\partial\mathcal{T}$ by

$$d_\kappa(x, y) = \kappa(x \wedge y), \quad x, y \in \partial\mathcal{T},$$

where $x \wedge y$ will denotes the deepest vertex through which x and y pass.

Conversely the Michon Theorem [27] establishes that any ultrametric Cantor set is isometric to the boundary of a rooted Cantorian tree endowed with a weight. Such a tree is unique if it is additionally required to be *reduced*, that is all of its vertices are branching. This tree is then called the Michon tree of the ultrametric Cantor set.

Any given tree can be reduced as follows without changing its boundary (up to isometry). Given \mathcal{T} , let $\mathcal{V}' \subset \mathcal{V}$ be the set of branching vertices which have infinitely many branching descendants (if \mathcal{T} is Cantorian, \mathcal{V}' is the set of all branching vertices). If $w \in \mathcal{V}'$, define its parent v in \mathcal{T}' to be the closest ancestor in \mathcal{T} which belongs to \mathcal{V}' . The new root \bullet' is either \bullet if the latter belongs to \mathcal{V}' or its closest descendant belonging to \mathcal{V}' . The vertex set \mathcal{V}' with the ancestor relation inherited from \mathcal{T} defines a tree \mathcal{T}' , which is reduced. If the original tree comes with a weight κ , then the restriction of κ to \mathcal{V}' gives a weight on the reduced tree. The two boundaries endowed with the induced ultrametries are then isometrically homeomorphic.

In the rest of this paper, each time an ultrametric Cantor set (X, d) is given, it is implied that the unique Michon tree which represents it (\mathcal{T}, κ) is given as well.

2.2. Assouad dimension. Assouad defined [3, 4] a metric dimension associated with a metric space. When the space is ultrametric, this dimension captures particularly well the extent to which the space can be embedded in an Euclidean space.

Definition 1. *Given $s > 0$, a metric space (X, d) is s -subhomogeneous if there exist $C > 0$ such that, for all $a > 0$ and all $b > 0$, and for all discrete subset $Y \subset X$,*

$$\left(\forall x \neq y \in Y, a \leq d(x, y) \leq b \right) \quad \Rightarrow \quad \#Y \leq C \left(\frac{b}{a} \right)^s.$$

Then, the metric dimension (or Assouad dimension) of X is defined by

$$\dim_A(X) = \inf \{s \in [0, +\infty]; X \text{ is } s\text{-subhomogeneous}\}$$

It is important to notice that the definition is uniform over the space X . Namely both constants C and s are independent on the location of the subset Y . It becomes clear that $\dim_A(X)$ is finite if and only if (X, d) has the doubling property.

Proposition 2. *The Assouad dimension is invariant under bi-Lipschitz homeomorphism.*

3. EMBEDDING THEOREMS

3.1. Known Results. For ultrametric Cantor sets the following Assouad Theorem holds

Theorem 4 (Assouad [3, 4]). *Let (X, d) be an ultrametric space. If $\dim_{\mathbb{A}}(X, d) < +\infty$, then (X, d) , is bi-Lipschitz embeddable in an Euclidean space.*

This result was later refined by Luukkainen and Movahedi-Lankarani [25]

Theorem 5. *Let (X, d) be an ultrametric space. If $\dim_{\mathbb{A}}(X) < n \in \mathbb{N}$, then X can be embedded in \mathbb{R}^n . Conversely, if X can be embedded in \mathbb{R}^n , then $\dim_{\mathbb{A}}(X) \leq n$.*

This last theorem provides an optimal dimension for embedding when the Assouad dimension of the space is not an integer. However, if $\dim_{\mathbb{A}}(X)$ is an integer, the theorem does not answer the question whether or not X is embeddable in \mathbb{R}^n . This is addressed by Luosto [24], and the answer is negative

Theorem 6 (Corollary 4.6 in [24]). *An ultrametric space (X, d) can be bi-Lipschitz embedded in \mathbb{R}^n if and only if $\dim_{\mathbb{A}}(X) < n$.*

3.2. Assouad Dimension and Michon Trees. The purpose of this Section is to give a description of the Assouad dimension for ultrametric Cantor sets in terms of its Michon tree representation.

Given a weighted Cantorian tree (\mathcal{T}, κ) , a vertex v of the tree and a number $0 < \delta < 1$, the *sub-tree of \mathcal{T} under v at resolution δ* (noted $\mathcal{T}(v, \delta)$) is defined as follows.

- Its vertices consist of all $w \in \mathcal{V}$ such that $w \preceq v$ and $\kappa(w) \geq \delta\kappa(v)$, as well as their children;
- The root is v ;
- Its edges are all edges of \mathcal{T} between vertices of the sub-tree.

Since \mathcal{T} is locally finite and the weight tends to 0 at infinity, this sub-tree is finite. By definition, a *leaf* (or minimal vertices for \succ) of this finite tree is a vertex w in the subtree satisfying $\kappa(w) < \delta\kappa(v)$. Hence leaves are the *closest descendants of v* satisfying this relation. Let $\mathcal{L}(\mathcal{T}, v, \delta)$ denote the set of all such leaves.

Proposition 3. *Let (X, d) be an ultrametric Cantor set with Michon tree (\mathcal{T}, κ) . Then X is s -subhomogeneous if and only if there exists a constant $C > 0$ such that for all vertex v in \mathcal{T} and all $0 < \delta < 1$,*

$$(1) \quad \#\mathcal{L}(\mathcal{T}, v, \delta) \leq C \delta^{-s}.$$

Remark 1. Since the number of leaves in $\mathcal{T}(v, \delta)$ is bounded from below by the number of children of v , it follows immediately that the Assouad dimension is infinite whenever the number of children per vertex is not uniformly bounded. \square

Proof. (i) Let (X, d) be s -subhomogeneous. For any vertex v and any $0 < \delta < 1$, let Y be the set made of exactly one point in each $[w]$ when w runs through the set of leaves of $\mathcal{T}(v, \delta)$. Since $Y \subset [v]$ it follows that $d(x, y) \leq \kappa(v)$ for $x, y \in Y$. Moreover, if $x \neq y$, then $d(x, y) = \kappa(u)$ for some vertex $u \in \mathcal{T}(v, \delta)$ which is an ancestor of the leaves in which x and y are chosen. Hence $d(x, y) \geq \delta\kappa(v)$. The s -homogeneity of (X, d) then gives the inequality (1).

(ii) Conversely, let (X, d) satisfies (1). Let $0 < a < b$, and let Y be a subset of X such that $a \leq d(x, y) \leq b$ for all $x, y \in Y$. In particular the diameter of Y is less than or equal to b . Hence there is a vertex v such that $Y \subset [v]$ with $\kappa(v) \leq b$. With each $x \in Y$ is associated the leaf w_x

of $\mathcal{T}(v, a/b)$ through which the path representing x is going. If $y \neq x$ is another point of Y , it follows that $w_y \neq w_x$. For otherwise $d(x, y) \leq \kappa(w_x) < \kappa(v) a/b \leq a$. Hence the cardinality of Y cannot be larger than the cardinality of $\mathcal{L}(\mathcal{T}, v, a/b)$. Thanks to eq. (1), it follows that (X, d) is s -homogeneous. \square

A criterion for checking finiteness of the Assouad dimension is given now. Remark that in conjunction with Theorem 4, this gives a criterion for embeddability into some \mathbb{R}^n .

Proposition 4. *Let (X, d) be an ultrametric Cantor set with Michon tree (\mathcal{T}, κ) . Then the Assouad dimension of X is finite if and only if the number of children per vertex is bounded (uniformly) and there are two constants $c \geq 1$ and $0 < \alpha < 1$ such that for all vertices $v \succeq w$,*

$$(2) \quad \frac{\kappa(w)}{\kappa(v)} \leq c \alpha^{|w|-|v|},$$

where $|v|$ (resp. $|w|$) is the depth of v (resp. w) in the tree.

Remark 2. This characterization in terms of the Michon tree can be compared with the doubling property for a metric space. In the present case, it corresponds to the condition that the cardinality of the sets $\mathcal{L}(\mathcal{T}, v, 1/2)$ is uniformly bounded. Assouad proved that the doubling property is equivalent to finite metric dimension. \square

Proof. (i) Let (2) hold. Then let M be the maximal number of children per vertex. For any $\delta > 0$, let $k - 1$ be the integer part of $\ln(\delta)/\ln(\alpha)$. If w is a descendant of v of depth $|v| + k$, then

$$\kappa(w) \leq c \kappa(v) \alpha^k \leq \kappa(v) \exp\left(\ln(\alpha) \frac{\ln(\delta)}{\ln(\alpha)}\right) = \kappa(v) \delta.$$

Therefore, elements in $\mathcal{L}(\mathcal{T}, v, \delta)$ are at depth at most $|v| + k$ and its cardinality is at most M^k . Consequently,

$$M^k \leq M \cdot M^{\ln(\delta)/\ln(\alpha)} = M \exp\left(\ln(\delta) \frac{\ln(M)}{\ln(\alpha)}\right) = M \delta^s,$$

with $s = \ln(M)/\ln(\alpha)$. Therefore, X is s -subhomogeneous, and the Assouad dimension is finite.

(ii) If the assumption represented by eq. (2) does not hold then $\forall C \geq 1, \forall 0 < \alpha < 1, \exists w \preceq v$ vertices in \mathcal{T} , such that

$$\frac{\kappa(w)}{\kappa(v)} > C \alpha^{|w|-|v|}.$$

In particular it implies that

$$(3) \quad \forall n \in \mathbb{N}, \exists v_n \succ w_n \text{ such that } |w_n| - |v_n| = n \text{ and } \frac{\kappa(w_n)}{\kappa(v)} > \frac{1}{2}.$$

For indeed, if not, let n be chosen such that for any vertex v and any descendant w of v which is n levels deeper, $\kappa(w)/\kappa(v) \leq 1/2$. Let now v and w be two arbitrary vertices such that $w \prec v$. By Euclidean division, $|w| - |v| = kn + r$ with $0 \leq r < n$. Then, let $v = w_0, w_1, \dots, w_k$ be vertices between v and w such that $|w_i| - |v| = in$. By assumption,

$$\frac{\kappa(w)}{\kappa(v)} \leq \frac{\kappa(w_k)}{\kappa(v)} \leq \prod_{i=1}^k \frac{\kappa(w_i)}{\kappa(w_{i-1})} \leq \frac{1}{2^k} \leq 2 \left(\frac{1}{2^{1/n}} \right)^{kn+r} = \left(\frac{1}{2^{1/n}} \right)^{|w|-|v|}.$$

So using the negation of equation (3), it was possible to establish a geometric rate of decay of the weights, which was assumed not to hold in the first place. Therefore, equation (3) holds.

(iii) If the assumption represented by eq. (2) does not hold then, using equation (3), there exists a sequence of vertices v_n such that $\mathcal{L}(\mathcal{T}, v_n, 1/2)$ contains w_n , which is n levels deeper than v_n . So, the subtree $\mathcal{T}(v_n, 1/2)$ contains all vertices between v_n and w_n , as well as all their children. Since each vertex is branching, it has at least two children. Therefore, $\mathcal{L}(\mathcal{T}, v_n, 1/2)$ contains at least n elements. This quantity tends to infinity as n does. Hence, using the characterization of s -subhomogeneity of Proposition 3, X cannot be s -subhomogeneous for any s , and its Assouad dimension is infinite. \square

3.3. Embedding Theorems and Michon trees. This section provides a proof of Theorem 5 for Cantor sets, using the Michon tree formalism.

Proposition 5. *Let (X, d) be an ultrametric Cantor set with Michon tree (\mathcal{T}, κ) . If (X, d) is bi-Lipschitz embeddable in \mathbb{R}^n , then the Assouad dimension of (X, d) is at most n .*

Proof. Suppose h is a bi-Lipschitz map $(X, d) \rightarrow \mathbb{R}^n$. Then, there are two constants $0 < m < M$ such that for any vertex v of \mathcal{T} ,

$$m\kappa(v) \leq \text{diam}(h([v])) \leq M\kappa(v).$$

The goal is to show that (X, d) is n -subhomogeneous, so that the Assouad dimension of (X, d) is at most equal to n . Let $0 < \delta < 1$, and let v be a vertex of \mathcal{T} . By definition of $\mathcal{L}(\mathcal{T}, v, \delta)$, given any distinct $w, w' \in \mathcal{L}(\mathcal{T}, v, \delta)$ and any $x \in w$ and $x' \in w'$, one has $d(x, x') = \kappa(w \wedge w') \geq \delta\kappa(v)$. Therefore,

$$d(h(x), h(x')) \geq m d(x, x') \geq m\delta\kappa(v).$$

Therefore, the Euclidean balls of center $h(x)$ and $h(x')$ and diameter $m\delta\kappa(v)$ are disjoint. It holds for any pair $w \neq w'$ of elements in $\mathcal{L}(\mathcal{T}, v, \delta)$. It means that $h([v])$, which is included in a ball of diameter $M\kappa(v)$ contains $\#\mathcal{L}(\mathcal{T}, v, \delta)$ disjoint balls of diameter $m\delta\kappa(v)$. By a comparing the volumes, one gets:

$$\#\mathcal{L}(\mathcal{T}, v, \delta)(m\delta\kappa(v))^n \leq (M\kappa(v))^n.$$

Rearranging the terms, it gives n -subhomogeneity with constant $C = (M/n)^n$. \square

Proposition 6. *Let (X, d) be a s -subhomogeneous space, with $s < n$. Then (X, d) is bi-Lipschitz embeddable in \mathbb{R}^n .*

The proof is adapted from the one in [19]. It relies on the construction of the δ -compressed tree ($0 < \delta < 1$) in an way analogous to the construction of telescoped diagrams in the self-similar case. The construction is given by the following lemma.

Lemma 1. *Let (\mathcal{T}, κ) be a weighted Cantorian reduced tree, which is s -subhomogeneous. Then, for all $0 < \delta < 1$, there exists $(\mathcal{T}_\delta, \kappa_\delta)$ a weighted Cantorian reduced tree called the δ -compression of (\mathcal{T}, κ) such that*

- (1) $\partial\mathcal{T}$ and $\partial\mathcal{T}_\delta$ are bi-Lipschitz homeomorphic (for the distances induced by their respective weight functions);
- (2) $\kappa_\delta(w) < \delta\kappa_\delta(v)$ for all parent-child pair (v, w) of vertices of \mathcal{T}_δ .

Furthermore, there exists a C (independent of δ), such that for all δ , the number of children per vertex in \mathcal{T}_δ is bounded uniformly by $C\delta^{-s}$.

Proof. (i) The δ -compression of \mathcal{T} is built inductively as follows: let δ be chosen in $(0, 1)$.

- (i) The root of \mathcal{T}_δ is \bullet , so that $(\mathcal{V}_\delta)_0 = \mathcal{V}_0$.
- (ii) $(\mathcal{V}_\delta)_{n+1} = \bigcup \{\mathcal{L}(\mathcal{T}, v, \delta) ; v \in (\mathcal{V}_\delta)_n\}$.
- (iii) $w \text{ in } (\mathcal{V}_\delta)_{n+1}$ is a child of $v \in (\mathcal{V}_\delta)_n$ in \mathcal{T}_δ if and only if w is a descendant of v in \mathcal{T} .
- (iv) The weight κ' on \mathcal{V}' is the restriction of κ .

The new tree \mathcal{T}_δ is always reduced if \mathcal{T} is, because a vertex v has more children in \mathcal{T}_δ than it has in \mathcal{T} , and therefore all vertices of \mathcal{T}_δ are branching. Also, by definition of $\mathcal{L}(\mathcal{T}, v, \delta)$, the weight satisfy condition (2) of the lemma.

Thanks to the s -subhomogeneity, there is $C > 0$ such that the number of children of a vertex in \mathcal{T}_δ is bounded by the cardinality of the sets $\mathcal{L}(\mathcal{T}, v, \delta)$ namely by $C\delta^{-s}$. It proves the last point. Another consequence is that \mathcal{T}_δ is locally finite, and therefore Cantorian.

(ii) The map $\phi : \partial\mathcal{T} \rightarrow \partial\mathcal{T}_\delta$ is defined as follows. Let $\gamma \in \partial\mathcal{T}$ be a path defined by the sequence of vertices (v_0, v_1, \dots) be the sequence of vertices through which it passes. Since $\lim_{k \rightarrow \infty} \kappa(v_k) = 0$, there is a sequence $(v_{k(i)})_{i \geq 0}$ such that $k(0) = 0$ and $k(i+1) = \min\{k > k(i) ; \kappa(v_{k(i+1)}) < \delta\kappa(v_{k(i)})\}$. Hence the $v_{k(i)}$ are vertices of \mathcal{T}_δ . Consequently the sequence of vertices $(v_{k(i)})_{i \geq 0}$ defines a unique path in \mathcal{T}_δ , which is set to be $\phi(\gamma)$ by definition. It is easy to see that ϕ is bijective, and that the image by ϕ of an open set is open. Therefore, ϕ^{-1} is continuous. Since \mathcal{T}_δ is locally finite (by the s -subhomogeneity assumption), $\partial\mathcal{T}_\delta$ is compact, so ϕ is a homeomorphism.

(iii) To see that ϕ is bi-Lipschitz, let $x, y \in \partial\mathcal{T}$. Let $v = x \wedge y$. Then $d(x, y) = \kappa(v)$. Let w be the closest ancestor of v which is also a vertex of \mathcal{T}_δ (possibly $v = w$; more precisely, v is a vertex of \mathcal{T}_δ if and only if $v = w$). Then by definition of \mathcal{T}_δ ,

$$(4) \quad \delta\kappa(w) \leq \kappa(v) \leq \kappa(w).$$

It is a quick check that $d_{\kappa_\delta}(\phi(x), \phi(y)) = \kappa_\delta(w)$. Together with equation (4), it proves that ϕ is bi-Lipschitz. \square

Proof of Proposition 6. Let (\mathcal{T}, κ) be the Michon tree associated with (X, d) . Let $0 < \delta < 1$, and $(\mathcal{T}_\delta, \kappa_\delta)$ be the δ -compression of the tree. The value of δ will be adjusted later. The goal is to first build a map $\partial\mathcal{T}_\delta \rightarrow \mathbb{R}^n$, and then show that this map is a bi-Lipschitz homeomorphism for an appropriate value of δ . Since (X, d) is bi-Lipschitz homeomorphic to $\partial\mathcal{T}_\delta$, it will prove the result.

(i) Let $M(\delta)$ be an upper bound for the number of children per vertex in \mathcal{T}_δ (there is a C such that for all δ , $M(\delta) \leq C\delta^{-s}$, with $s < n$). Then, define a function $g : \mathcal{V}_\delta \rightarrow \{1, \dots, M(\delta)\}$ such that its restriction to the set of children of v be one-to-one for all vertex $v \in \mathcal{V}_\delta$. Such a map is simply a numbering of children of each vertex. Any point $x \in \partial\mathcal{T}_\delta$ will be represented by the sequence $x = (v_k)_{k \geq 0}$ of its vertices in \mathcal{T}_δ , with v_k a vertex of depth k , for all k . In particular (by definition of \mathcal{T}_δ), $\kappa_\delta(v_k) < \delta^k$. Then let $\phi : \partial\mathcal{T}_\delta \rightarrow \mathbb{R}^n$ be defined by $\phi(x) = (\phi_r(x))_{r=1}^n$ with

$$\phi_r(x) = \sum_{j=0}^{\infty} g(v_{jn+r})\kappa_\delta(v_{jn+r-1}), \quad 1 \leq r \leq n.$$

Since $g(v_k) \leq M(\delta)$ and $\kappa_\delta(v_k) < \delta^k$ it follows that the series converges absolutely and uniformly with respect to r and x .

(ii) Let now $x, y \in \partial\mathcal{T}$ with $x \neq y$ with $x = (v_k)_{k \geq 0}$ and $y = (w_k)_{k \geq 0}$. Then let j_0, r_0 be natural integers such that $|x \wedge y| = j_0 n + r_0 - 1$ and $1 \leq r_0 \leq n$. It follows that the paths x and y share the same vertices until the generation $|x \wedge y|$ and they split at the next generation. Therefore (with the notation that $\chi(r \geq r_0)$ is 1 if $r \geq r_0$ and 0 otherwise),

$$\begin{aligned} \phi_r(x) - \phi_r(y) &= \sum_{j=j_0+1}^{\infty} (g(v_{jn+r})\kappa(v_{jn+r-1}) - g(w_{jn+r})\kappa(w_{jn+r-1})) \\ &\quad + \chi(r \geq r_0)(g(v_{j_0n+r})\kappa(v_{j_0n+r-1}) - g(w_{j_0n+r})\kappa(w_{j_0n+r-1})) \end{aligned}$$

It should be remarked that

$$-g(w_{jn+r})\kappa(w_{jn+r-1}) \leq g(v_{jn+r})\kappa(v_{jn+r-1}) - g(w_{jn+r})\kappa(w_{jn+r-1}) \leq g(v_{jn+r})\kappa(v_{jn+r-1}).$$

Furthermore, both v_{jn+r-1} and w_{jn+r-1} are descendants of $x \wedge y$, sitting $n(j - j_0) + (r - r_0)$ levels deeper than $x \wedge y$. Therefore, by construction of \mathcal{T}_δ ,

$$\kappa(v_{jn+r-1}) < \delta^{n(j-j_0)+(r-r_0)} \kappa(x \wedge y), \quad \text{and similarly for } w_{jn+r-1}.$$

It leads to the inequality

$$|g(v_{jn+r})\kappa(v_{jn+r-1}) - g(w_{jn+r})\kappa(w_{jn+r-1})| \leq M(\delta) \kappa(x \wedge y) \delta^{n(j-j_0)+(r-r_0)}$$

Therefore

$$(5) \quad |\phi_r(x) - \phi_r(y)| \leq \frac{M(\delta)}{1 - \delta^n} \kappa(x \wedge y).$$

Since $\kappa(x \wedge y) = d_\kappa(x, y)$, this map is Lipschitz continuous. On the other hand, $\|\phi(x) - \phi(y)\| \geq |\phi_r(x) - \phi_r(y)|$ for all $1 \leq r \leq n$. In particular, using again eq. (5) and since $v_{j_0n+r_0} \neq w_{j_0n+r_0}$, it follows that

$$\|\phi(x) - \phi(y)\| \geq |\phi_{r_0}(x) - \phi_{r_0}(y)| \geq \kappa(x \wedge y) - \frac{\delta^n M(\delta)}{1 - \delta^n} \kappa(x \wedge y).$$

Now, since $M(\delta)$ grows at most like $C\delta^{-s}$, with $s < n$, it follows that for a small enough value of δ , this map is actually bi-Lipschitz (in particular, it is one-to-one). \square

4. S -ADIC SYSTEMS

4.1. Definitions. This section is devoted to the basic definitions for S -adic systems. The presentation uses the notations of Durand [11, 12]. Let A be a finite alphabet. Note A^* the set of all finite words with letters in A . For $w \in A^*$, $|w|$ denotes its length, namely the number of letters it contains. Consider S a *finite* set of morphisms

$$\sigma \in S : A(\sigma) \longrightarrow A^*, \quad \text{with } A(\sigma) \subset A.$$

A S -adic system is a sequence $(\sigma_n : A_{n+1} \rightarrow (A_n)^*)_{n \in \mathbb{N}} \in S^{\mathbb{N}}$, such that the morphisms are composable. It will be assumed that for all n , every letter in A_n appears in a word $\sigma_n(b)$ for some $b \in A_{n+1}$. For $m > n$ the following notation will be used

$$\sigma_{n,m} = \sigma_n \circ \dots \circ \sigma_{m-1} : A_m \rightarrow A_n.$$

It satisfies $\sigma_{n,m} \circ \sigma_{m,k} = \sigma_{n,k}$. A S -adic system is *primitive* if there exists some $s_0 > 0$ such that for all $r \in \mathbb{N}$, for all $a \in A_{r+s_0}$ and all $b \in A_r$, the letter b appears in the word $\sigma_{r,r+s_0}(a)$. It is called *proper* if there exist two letters l and r in A such that for all $\sigma \in S$ and all $a \in A(\sigma)$, $\sigma(a)$ begins by the letter l and ends by r . Furthermore, it will be assumed from this point on, that

$$(6) \quad \lim_{n \rightarrow +\infty} \min_{c \in A_n} |\sigma_{1,n}(c)| = +\infty,$$

so that words of arbitrary long length are obtained from iterating the substitutions on a single letter.

Given a proper S -adic system, there is a way to associate a subshift $\Xi \subset A^{\mathbb{Z}}$ (it is then called a S -adic subshift). Let T be the shift operator on $A^{\mathbb{Z}}$

$$T(\dots x_{-1} \cdot x_0 x_1 \dots) = \dots x_{-1} x_0 \cdot x_1 \dots$$

It is straightforward to extend the morphisms of S to $A^{\mathbb{Z}}$ by concatenation

$$\sigma(\dots x_{-1} \cdot x_0 x_1 \dots) = \dots \sigma(x_{-1}) \cdot \sigma(x_0) \sigma(x_1) \dots$$

It ought to be remarked that, by properness, for all n , $\sigma_{1,n}(l)$ is a prefix of $\sigma_{1,n+1}(l)$. Similarly, $\sigma_{1,n}(r)$ is a suffix of $\sigma_{1,n+1}(r)$. Therefore, an element of $A^{\mathbb{Z}}$ can be defined by

$$x = \left(\lim_{n \rightarrow +\infty} \sigma_{1,n}(r) \right) \cdot \left(\lim_{n \rightarrow +\infty} \sigma_{1,n}(l) \right),$$

where the dot separates the x_i with $i \geq 0$ on the right from the ones with $i < 0$ on the left. The subshift associated with the S -adic system is the closure in $A^{\mathbb{Z}}$ (endowed with the product topology) of the orbit of x under the shift. This subshift, Ξ , is endowed with the combinatorial distance d , namely two sequences have distance $(n+1)^{-1}$ whenever they coincide on a string of radius n around the origin and do not coincide beyond. If the S -adic system is primitive, (Ξ, T) is minimal (see Durand [11, Lemma 7]).

Definition 2. A word $x \in A^{\mathbb{Z}}$ is called *linearly recurrent* or *linearly repetitive* (LR for short) if there is a constant K such that for all $n \in \mathbb{N}$, for all subwords w and w' of x respectively of length n and Kn , then w occurs in w' as a subword.

In other terms, in a LR word, each finite subword repeats infinitely often and within a distance which varies linearly with its size. It is easy to see that a subshift generated by an LR word is minimal and that every elements of the subshift is LR with the same constant. In this case, the subshift itself is called *linearly recurrent*. The following is due to Durand.

Theorem 7 (Durand [12], Proposition 1.1). *A subshift is S -adic primitive and proper if and only if it is linearly recurrent.*

The periodic case is trivial. In the non-periodic case, Durand's construction of the S -adic system associated with a linearly recurrent subshift involves return words. The S -adic system he builds in this case has the unique decomposition property (defined below). This property will be used later on (see [12, Section 4] for the construction, and [13, Definition 9, Lemma 17] for the properties of return words and codes).

Definition 3. Let $(\sigma_n)_{n \in \mathbb{N}}$ be a S -adic system. It is said to have *unique decomposition property* if, given any element x in the subshift associated with Ξ , there is a unique decomposition of x as a concatenation

$$x = \dots \sigma_1(a_{-1}) \underline{\sigma_1(a_0)} \sigma_1(a_1) \dots,$$

where the index 0 of x is in the underlined word $\sigma_1(a_0)$.

4.2. Proof of Theorem 1. The periodic case is trivial (in this case, Ξ is finite), and so it may be assumed that the subshift is aperiodic. Using the theorem of Durand cited above, it is enough to prove the following

Proposition 7. *Let Ξ be a primitive proper S -adic subshift with unique decomposition property and endowed with the combinatorial metric. Then, it is f -embeddable.*

It will be convenient to describe an S -adic subshift in terms of a Bratteli diagram (see for example [13]). A weight on the Bratteli diagram permits to define an ultrametric on the set of its infinite paths. It will be necessary to prove that:

- the ultrametric Cantor set associated with this Bratteli diagram is bi-Lipschitz homeomorphic to the subshift with the combinatorial metric;
- the weight on the Bratteli diagram satisfies the decay rate of Proposition 4 so that the Cantor set associated with the diagram is embeddable.

The following result is needed.

Lemma 2 (Durand [11], Lemma 8). *If the S -adic system generated by $(\sigma_n)_{n \in \mathbb{N}}$ is primitive with constant s_0 , there exists a constant K such that for all integers r, s , with $s - r \geq s_0$ and for all b, c in A_{s+1}*

$$\frac{|\sigma_{r,s+1}(b)|}{|\sigma_{r,s+1}(c)|} \leq K.$$

Let $(\sigma_n : A_{n+1} \rightarrow A_n)$ be a primitive S -adic proper system with unique decomposition property and let l and r the letters associated with the properness. The Bratteli diagram is defined as follows

- For all $n \in \mathbb{N}$, the set of vertices \mathcal{V}_n is in bijection with the alphabet A_n : for each $a \in A_n$, there is a $v_a \in \mathcal{V}_n$.
- For all $n \in \mathbb{N}$, an edge in \mathcal{E}_n is a triple $e = (v_a, k, v_b)$ where $v_a \in \mathcal{V}_n$, $v_b \in \mathcal{V}_{n+1}$ and $k \in \mathbb{N}$ such that the letter a occurs in the word $\sigma_n(b)$ in position $k + 1$. Hence, the number of edges from v_a to v_b is equal to the number of times the letter a appears in $\sigma_n(b)$.
- If $e = (v_a, k, v_b) \in \mathcal{E}_n$ its *source* is $s(e) = v_a$, its *range* is $r(e) = v_b$ and its *label* is $\ell(e) = k$.

Example 1. If $\sigma_n(b) = labcar$, then there are two edges e_1, e_2 from v_a to v_b : one corresponds to the second letter and the other to the fifth letter of the word above. Their respective labels are 1 and 4. \square

Definition 4. *A path on the Bratteli diagram is a sequence of edges $(e_k)_{i \leq k < j}$ (with $1 \leq i < j \leq +\infty$), such that $e_k \in \mathcal{E}_k$ for all k and $r(e_k) = s(e_{k+1})$. Let Π_n denote the set of paths with $i = 1$, $j = n$ (simply called “paths of length n ”). Let Π denote the union of all Π_n and let Π_∞ be the set of paths of infinite length ($i = 1$, $j = +\infty$).*

Because the σ_n 's are taken from a finite set of substitutions S the cardinality of the sets \mathcal{E}_n and \mathcal{V}_n is uniformly bounded in n . Let

$$w_n = \left(\min\{|\sigma_{1,n}(a)| ; a \in A_n\} \right)^{-1}.$$

It is a decreasing sequence which tends to 0 as $n \rightarrow \infty$. The weight of a finite path γ will be defined by w_{n+1} whenever γ has length n . This leads to a metric on Π_∞ defined by

$$d_w(x, y) = w_{n+1}, \quad \text{where } n \text{ is the length of the longest common prefix of } x, y.$$

Proposition 8. *The subshift (Ξ, d) , endowed with the combinatorial metric, and (Π_∞, d_w) are homeomorphic through a bi-Lipschitz homeomorphism.*

Proof. (i) **Constructing a map $\Xi \rightarrow \Pi_\infty$.** Let $x = \dots x_{-1} \cdot x_0 x_1 \dots$ in Ξ . Let $v_1 \in \mathcal{V}_1$ be the vertex corresponding to the letter x_0 . By the unique decomposition property, x can be written in a unique way as the concatenation

$$x = \dots \sigma(x'_{-1}) \underline{\sigma(x'_0)} \sigma(x'_1) \dots ,$$

with $(x'_i)_{i \in \mathbb{Z}} \in (A_2)^\mathbb{Z}$ and such that the letter of x of index 0 is in the underlined word $\sigma(x'_0)$. Therefore, if x' is the word $x' = (x'_i)_{i \in \mathbb{Z}}$

$$(7) \quad x = T^k \sigma(x'),$$

for some $0 \leq k < |\sigma(x'_0)|$. This label k corresponds to an occurrence of the letter x_0 in the word $\sigma(x'_0)$. It defines an edge $e = (v_1, k, v_2)$, where $v_2 \in \mathcal{V}_2$ is the vertex corresponding to x'_0 . Iterating this process of “de-substitution” leads to construct a sequence of words $x', x'', \dots, x^{(n)} \dots$ and a corresponding sequence of edges e_1, e_2, \dots, e_n . In particular $\psi(x) = (e_1, e_2, \dots)$ defines a map $\Xi \rightarrow \Pi_\infty$.

(ii) **ψ is a bijection.** The map ψ has an inverse $\phi : \Pi_\infty \rightarrow \Xi$ which will be built explicitly. Let $\gamma = (e_1, e_2, \dots)$ be a path going through the vertices v_1, v_2, \dots . Let ϵ be a symbol not in A (the “empty” symbol). Now, given a finite word $w = (w_0, \dots, w_{l-1}) \in A^*$, Define then $\bar{w} \in (A \cup \{\epsilon\})^\mathbb{Z}$ as the sequence

$$\bar{w} = \dots \epsilon \epsilon \epsilon \cdot w \epsilon \epsilon \epsilon \dots$$

This is a way of seeing a word in A^* as a (partially defined) element of $A^\mathbb{Z}$. For fixed n , define the sequence

$$(8) \quad \phi_n(\gamma) = T^{\ell(e_1)} \circ \sigma_1 \circ T^{\ell(e_2)} \circ \dots \circ \sigma_{n-2} \circ T^{\ell(e_{n-1})+1} \left(\overline{r \sigma_{n-1}(a_n) l} \right) ,$$

where $\ell(e)$ denotes the label of the edge e . This sequence can be seen as an element of $(A \cup \{\epsilon\})^\mathbb{Z}$, namely as:

$$\dots \epsilon \epsilon \sigma_{1,n-1}(r) \underline{\sigma_{1,n}(a_n)} \sigma_{1,n-1}(l) \epsilon \epsilon \dots ,$$

where the letter of index 0 occurs in the underlined word. Its exact position is determined by the labels of the edges. This implies

$$(9) \quad [\phi_n(\gamma)]_i \neq \epsilon \quad \text{for} \quad -|\sigma_{1,n-1}(r)| \leq i \leq |\sigma_{1,n-1}(l)|.$$

By hypothesis, $\lim_{n \rightarrow +\infty} |\sigma_{1,n-1}(b)| = +\infty$ and this is true, in particular, whenever $b \in \{r, l\}$. Therefore, the limit as n tends to infinity of $\phi_n(\gamma)$ is an element of $A^\mathbb{Z}$, which is noted $\phi(x)$. Using the definitions of ϕ and ψ (see eq. (7) & (8)), it is straightforward that $\psi \circ \phi = \text{id}_{\Pi_\infty}$. Therefore, ψ is onto and ϕ is one-to-one.

Conversely, let $x \in \Xi$, and $\gamma = \psi(x)$. It will be shown that $\phi(\gamma) = x$. For all n , x has a unique decomposition

$$x = \dots \sigma_{1,n}(b_{-1}) \underline{\sigma_{1,n}(b_0)} \sigma_{1,n}(b_1) \dots ,$$

with $b_i \in A_n$. So by definition of ψ and ϕ ,

$$(10) \quad \phi_n(\gamma) = \dots \epsilon \sigma_{1,n-1}(r) \underline{\sigma_{1,n}(b_0)} \sigma_{1,n-1}(l) \epsilon \dots,$$

and the two words coincide on the word $\sigma_{1,n}(b_0)$ (it appears at the same position). Furthermore, by properness of the S -adic system, $\sigma_n(b_{-1})$ ends with the letter r while $\sigma_n(b_1)$ begins with the letter l . Therefore, $\sigma_{1,n}(b_{-1})$ has $\sigma_{1,n-1}(r)$ as a suffix and $\sigma_{1,n}(b_1)$ has $\sigma_{1,n-1}(l)$ as a prefix. So for all $i \in \mathbb{Z}$ such that $[\phi_n(\gamma)]_i \neq \epsilon$, one has $[\phi_n(\gamma)]_i = x_i$. Taking a limit and using Equation (6), $\phi(\gamma)$ and x agree everywhere. Therefore, ϕ and ψ are inverse bijections of each other.

(iii) **ϕ is bi-Lipschitz.** Let $\gamma, \gamma' \in \Pi_\infty$ be such that the n first edges of γ and γ' coincide. Then, by definition, $\phi_n(\gamma) = \phi_n(\gamma')$. Thus $\phi(\gamma)$ and $\phi(\gamma')$ coincide for all indices i satisfying $-|\sigma_{1,n-1}(r)| \leq i \leq |\sigma_{1,n-1}(l)|$. Using Lemma 2 and the definition of w_n leads to

$$\forall b \in A_{n-1}, \quad (w_{n-1})^{-1} \leq |\sigma_{1,n-1}(b)| \leq K(w_{n-1})^{-1}.$$

Since S is finite, $C = \max_{\sigma \in S, a \in A(\sigma)} |\sigma(a)|$ is well defined. Then, a word of the form $\sigma_{1,n}(b) = \sigma_{1,n-1} \circ \sigma_{n-1}(b)$ is at most C times longer than the longest word of the form $\sigma_{1,n-1}(c)$. That is, using again Lemma 2, $(w_n)^{-1} \leq CK(w_{n-1})^{-1}$. If $[\phi(\gamma)]_i = [\phi(\gamma')]_i$ for $|i| \leq K(w_{n-1})^{-1}$, then in particular this inequality holds for $|i| \leq (w_n)^{-1}/C$. So

$$d(\gamma, \gamma') \leq w_n \quad \Rightarrow \quad d(\phi(\gamma), \phi(\gamma')) \leq Cw_n.$$

Conversely, let γ and γ' coincide up to edge n , but differing on their $(n+1)$ -th one. Then, by definition of ϕ

$$\begin{aligned} \phi(\gamma) &= \dots \sigma_{1,n+1}(b_{-1}) \underline{\sigma_{1,n+1}(b_0)} \sigma_{1,n+1}(b_1) \dots \\ \phi(\gamma') &= \dots \sigma_{1,n+1}(b'_{-1}) \underline{\sigma_{1,n+1}(b'_0)} \sigma_{1,n+1}(b'_1) \dots \end{aligned}$$

where $b_0 \neq b'_0$ and the letter of index 0 belongs to the respective underlined words. By the unique decomposition property, $\phi(\gamma)$ and $\phi(\gamma')$ have to differ at an index i satisfying

$$-\max\{|\sigma_{1,n+1}(b_0)|, |\sigma_{1,n+1}(b'_0)|\} \leq i \leq \max\{|\sigma_{1,n+1}(b_0)|, |\sigma_{1,n+1}(b'_0)|\},$$

that is for

$$-K(w_n)^{-1} \leq i \leq K(w_n)^{-1}.$$

This proves that ϕ^{-1} is Lipschitz. □

Proof of Proposition 8. In order to prove that (Ξ, d) is f -embeddable, it is sufficient to prove that (Π_∞, d_w) is embeddable. It suffices to remark that the corresponding Michon tree has the finite paths in Π as vertices and the parent map is induced by deletion of the last edge. By finiteness of S , all paths in Π_n have a bounded number of extensions to paths of Π_{n+1} and this bound is independent of n . Hence the Michon graph has a bounded number of children per vertex. The only thing left to show is that the sequence of weights $(w_n)_{n \in \mathbb{N}}$ on the Bratteli diagram is bounded above by a decreasing geometric sequence.

First, remark that for all n , the alphabet A_n has at least two elements. Indeed, if one of the A_n had only one element, the subshift Ξ would be periodic (this case was ruled out). Using primitivity, for all letter $c \in A_{s_0+1}$, $|\sigma_{s_0}(c)| \geq 2$. By iteration, for all k and all $c \in A_{k s_0+1}$

$$|\sigma_{ks_0}(c)| \geq 2^k.$$

So $w_{ks_0} \leq 2^{-k}$. Since $(w_n)_{n \in \mathbb{N}}$ is decreasing, $w_n \leq C\lambda^n$ with $\lambda = 2^{-1/s_0} < 1$, for some constant C . Thanks to Proposition 4 it is proved that (Ξ, d) has finite Assouad dimension, and is therefore embeddable.

4.3. An possible analogue result for tilings. It is a natural question to ask whether linearly repetitive tilings of \mathbb{R}^n are f -embeddable. It seems reasonable to make the following conjecture (see for example [1] for the definitions of the objects).

Conjecture 1. *Let Ξ be the transversal of a tiling space of linearly repetitive tilings of \mathbb{R}^d , together with the usual tiling metric. Then Ξ is f -embeddable.*

The proof for subshifts is based on the fact that linearly repetitive subshifts have a good representation by Bratteli diagrams. While they are not self-similar, there are only a finite number of substitutions involved, which allows to generalize the methods of [19]. In a recent article, Aliste and Coronel [1] provide a good description of linearly repetitive tiling spaces of any dimension as the set of paths on a Bratteli diagram. Given the transversal of a linearly repetitive tiling space, it is possible to represent it by the set of paths on a Bratteli diagram such that the number of vertices and edges in each \mathcal{V}_n or \mathcal{E}_n is bounded uniformly in n . There are also have estimates for the size of the patches associated with paths of size n . These estimates are analogues of Durand’s Lemma 2 and allow to define reasonable weights on the Bratteli diagram. In the end, the path space of this Bratteli diagram, endowed with the distance defined by the weight, is f -embeddable. However, there is no proof yet that the homeomorphism between this paths space and the tiling space is bi-Lipschitz.

In Proposition 8, the bi-Lipschitz character of ϕ is proved through using equation (10). This requires an estimate on the length of $\sigma_{1,n}(b_0)$, of $\sigma_{1,n-1}(r)$ and of $\sigma_{1,n-1}(l)$. This is a quantitative version of a property known as “forcing the border”. It seems that Aliste–Coronel’s construction does satisfy a similar quantitative border forcing property, but this fact was not highlighted as such in their paper (as it seems to be a byproduct of their construction and not an essential feature of the result they prove). Establishing it would require a thorough reworking of their already sophisticated proof. This justifies the fact that, even though this conjecture can be stated with some confidence, it is a conjecture and not as a result.

5. STURMIAN SEQUENCES

5.1. Definitions and notations. First, a few facts need to be introduced about Sturmian sequences and their coding. This section follows in part the presentation in [2], and proof for some of the results cited below can be found there.

Definition 5. *Given a sequence $x = (x_n)_{n \in \mathbb{Z}} \in A^{\mathbb{Z}}$ on the finite alphabet A , define its language:*

$$L_n(x) = \{\text{finite words } a_1 \dots a_n \text{ which appear in } x\}, \text{ and } L(x) = \bigcup_{n \in \mathbb{N}} L_n(x).$$

Definition 6. *Given $x \in A^{\mathbb{Z}}$, its complexity function p_x is defined by*

$$\forall n \in \mathbb{N}, \quad p_x(n) = \text{Card}(L_n).$$

When there is no risk of ambiguity, $p_x(n)$ is just noted $p(n)$. It is known that if $p_x(n) \leq n$ for some n , then the word x is periodic.

Definition 7. A sequence is called Sturmian if its complexity function satisfies $p(n) = n + 1$ and if it is not eventually periodic, namely if it doesn't have a one-sided periodic infinite prefix or suffix).

From the definition, a Sturmian sequence is a sequence on two letters (since $p(1) = 2$). These two letters are noted 0 and 1. Note that the “not eventually periodic” condition is here to rule out degenerate cases like $\dots 111000\dots$ or $\dots 0001000\dots$.

Proposition 9. The frequency of the letter 1 in a Sturmian sequence x

$$\text{freq}_x(1) := \lim_{n \rightarrow +\infty} \frac{\text{Card}\{k \in [-n, n] ; x_k = 1\}}{2n + 1}$$

is well defined, and is an irrational number, noted $\alpha \in]0, 1[$.

Definition 8. The Sturmian sequence x is of type zero if $\text{freq}_x(1) < 1/2$, and of type one if $\text{freq}_x(1) > 1/2$.

It is easy to see that in a Sturmian sequence of type 1, the words 10, 01 and 11 may appear, but not the word 00. The same statement holds for sequences of type 0, just exchanging the letters 0 and 1.

Given a Sturmian sequence x , there is naturally a subshift of $\{0, 1\}^{\mathbb{Z}}$ associated with it. It is by definition the set of all sequences y such that $L(y) = L(x)$. It can also be defined as the closure of the orbit of x in $A^{\mathbb{Z}}$ (for the product topology). The two definitions are equivalent. It is of course shift-invariant (hence the name “subshift”), and it is well known that it is minimal. If x has frequency of ones equal to α , so do all the elements of its subshift. In particular, it makes sense to write that a Sturmian subshift is of type 0 or of type 1. Conversely, the subshift associated with x is exactly the set of all Sturmian sequences with the same frequency as x . Therefore, it makes sense to denote $\Xi(\alpha)$ the subshift associated with $\alpha \in (0, 1)$. Sturmian sequences can be *recoded* using the following substitutions. Define:

$$\sigma_0 : \begin{cases} 0 & \mapsto 0 \\ 1 & \mapsto 10 \end{cases} \quad \text{and} \quad \sigma_1 : \begin{cases} 0 & \mapsto 01 \\ 1 & \mapsto 1. \end{cases}$$

Proposition 10. For any Sturmian sequence x of type 0, there is a Sturmian sequence x' such that either $x = \sigma_0(x')$ or $x = T\sigma_0(x')$, where T is the shift operator.

For any Sturmian sequence x of type 1, there is a Sturmian sequence x' such that either $x = \sigma_1(x')$ or $x = T\sigma_1(x')$.

Let Φ denote the recoding map $x \mapsto x'$. It is worth noting that if x, y are two elements of the same Sturmian subshift, then $\Phi(x)$ and $\Phi(y)$ belong to the same subshift. Non-periodicity of Sturmian sequences implies that for any Sturmian sequence y of type 0 (resp. 1), there is a k such that $\Phi^k(y)$ is of type 1 (resp. 0).

Definition 9. Let x be a Sturmian sequence, and $(\Phi^n(x))_{n \in \mathbb{N}}$ be the sequence of recoded Sturmian sequences. By definition, for all n ,

$$\sigma_0^{b_1} \circ \sigma_1^{b_2} \circ \dots \circ \sigma_0^{b_{2n+1}}(\Phi^{2n+1}(x)) \quad \text{and} \quad \sigma_0^{b_1} \circ \sigma_1^{b_2} \circ \dots \circ \sigma_1^{b_{2n}}(\Phi^{2n}(x))$$

are in the same orbit as x . All the b_n are positive, except maybe $b_0 = 0$. The sequence $(b_n)_{n \in \mathbb{N}}$ is called the multiplicative coding of the Sturmian sequence x .

All Sturmian sequences in a same subshift have the same multiplicative coding, and an acceptable multiplicative coding determines uniquely a Sturmian subshift.

5.2. Partial fraction decomposition and multiplicative coding. The properties of a Sturmian subshift are closely related to the partial fraction decomposition of the number α (which is the frequency of ones in the subshift).

Let $\alpha \in \mathbb{R}$ be irrational. Then α can be written uniquely $\alpha = a_0 + \alpha_0$, with $a_0 \in \mathbb{Z}$ and $\alpha_0 \in (0, 1)$. The Gauss map $G : [0, 1] \rightarrow [0, 1]$, applied to α_0 , generates the continuous fraction expansion

$$G(\alpha) = \frac{1}{\alpha} - a(\alpha), \quad a(\alpha) = \left[\frac{1}{\alpha} \right],$$

where $[x]$ denotes the *integer part* of x , namely the largest integer smaller than or equal to x . Hence

$$\alpha = a_0 + \frac{1}{a_1 + \alpha_1}, \quad a_1 = a(\alpha), \quad \alpha_1 = G(\alpha).$$

Iterating this formula gives rise to the continuous fraction expansion

$$\alpha = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots + \frac{1}{a_n + \alpha_n}}}}, \quad \alpha_{n+1} = G(\alpha_n), \quad a_n = a(\alpha_{n-1}) \quad n \geq 1.$$

The standard notation is

$$\alpha = [a_0; a_1, a_2, \dots, a_n, \dots],$$

and the a_n 's are called the *partial quotients* of α .

Definition 10. A number α has bounded type whenever, the sequence of its partial quotients is bounded.

Having bounded type is an exceptional property. This is a theorem of Khintchine [22]. See also Levy [23].

Theorem 8. For almost every $\alpha \in [0, 1]$ the sequence $(a_n)_{n \in \mathbb{N}_*}$ of partial quotients of α is unbounded.

One famous example of such a typical number is $e = 2, 71828 \dots$ the continued fraction of which was computed by Leonhard Euler in 1737 namely

$$e = [2, 1, 2, 1, 1, 4, 1, 1, 6, \dots, 1, 1, 2l, \dots] \quad a_{3l-1} = 2l, \quad a_{3l-2} = a_{3l} = 1, \quad l \geq 1.$$

Many properties of a Sturmian subshift are determined by the arithmetic properties of the number α associated with it. The following theorem is proved in Hedlund and Morse's seminal paper.

Theorem 9 (see [28]). Consider a Sturmian subshift $\Xi \subset \{0, 1\}^{\mathbb{Z}}$, of parameter α . This subshift is linearly repetitive (see Definition 2) if and only if α has bounded type.

An immediate consequence of this theorem and of Theorem 1 is the following result.

Proposition 11. *Let Ξ be a Sturmian subshift with parameter α , endowed with the combinatorial metric. If α has bounded type, then Ξ is f -embeddable.*

This result needs a converse statement: if α has not bounded type, then the associated Sturmian subshift is not f -embeddable. This will be proved in Section 5.3.

One way to understand the deep links between the combinatoric properties of a Sturmian subshift and the arithmetic properties of its parameter α is the following. A Sturmian sequence x can be seen as the coding of an orbit of the rotation of angle α on the circle: there is $\{I_0, I_1\}$ a partition of the circle, and s a point in the circle such that $x_n = 0$ if and only if $(s + \alpha \bmod 1) \in I_0$. The rotation on the circle can be related to the continued fraction decomposition of α on the one hand, and on the multiplicative coding of x on the other hand, to get to following result.

Theorem 10 ((see [2]). *Let x be a Sturmian sequence, and $\alpha = [0; a_1, a_2, \dots]$ be the frequency of 1 in x . Then the multiplicative coding of x is given by the partial quotients of $(1 - \alpha)/\alpha = \alpha^{-1} - 1$.*

It is straightforward that $(1 - \alpha)/\alpha = [a_1 - 1; a_2, a_3, \dots]$. In particular, α has bounded type if and only if the sequence of coefficients of the multiplicative coding is bounded.

5.3. Non embeddability of certain Sturmian subshifts. This section is devoted to the proof of the following result.

Proposition 12. *Consider a Sturmian sequence x , with associated subshift Ξ and associated parameter α . If α has unbounded partial quotients, then Ξ (with the combinatorial metric) is not embeddable in a finite dimensional space.*

A Sturmian subshift Ξ is well described combinatorially by the (bilateral) tree of words of its elements.

Definition 11. *Given a Sturmian subshift Ξ , its un-reduced tree of words is defined as follows.*

- For all $n \geq 0$, define a sequence of refining partitions of Ξ by:

$$\mathcal{P}_n = \{[y_{-n} \dots y_n] ; y \in \Xi\},$$

where $[y_{-n} \dots y_n]$ is the cylinder set of all words in Ξ which coincide with y on indices $-n \leq i \leq n$. By convention, $\mathcal{P}_{-1} := \{\Xi\}$.

- The set of vertices of the tree is in bijection with the disjoint union of all the \mathcal{P}_n . If $U \in \mathcal{P}_n$, the associated vertex is noted v_U .
- The ancestor relation is induced by inclusion: $v_U \preceq v_V$ if and only if $U \subseteq V$.
- The weight of the vertex v_U is $1/(n + 2)$ if $U \in \mathcal{P}_n$ ($n \geq -1$).

The (reduced) tree of words is obtained from this tree by the reduction process defined in section 2.1.

One remark about this tree: if $X \in \mathcal{P}_n$, then $\text{diam}(X) \leq n^{-1}$, for the combinatorial distance, with equality if and only if X is the non-trivial union of two distinct elements of \mathcal{P}_{n+1} . Note that the vertices v_X of such clopen sets X are exactly the vertices with two children: these are precisely the ones which are not dropped by the reduction process.

This leads to the following proposition.

Proposition 13. *The boundary of the reduced tree of words associated with a Sturmian subshift Ξ is bi-Lipschitz homeomorphic to the subshift (endowed with the combinatorial metric).*

Proof. The proof is almost tautological: given an infinite path in the tree, say γ , the sequence of vertices $(v_{X_n})_{n \in \mathbb{N}}$ defines a decreasing sequence of compact sets X_n , the diameter of which tends to zero. Therefore, its intersection is not empty and consists of a single element $\{x\}$. Define $\phi(\gamma) = x$.

Conversely, given $x \in \Xi$, there is a unique decreasing sequence of sets $X_n \in \mathcal{P}_n$, such that for all n , $x \in X_n$ (explicitly: $X_n = [x_{-n} \dots x_n]$). Then it defines an infinite path $\gamma \in \partial\mathcal{T}$. Clearly, γ is the unique pre-image of x by ϕ .

The fact that ϕ is bi-Lipschitz (and in particular, it is a homeomorphism) results from the remark above, on the diameters of the elements of \mathcal{P}_n . \square

Lemma 3. *Let y be a Sturmian sequence, assume that $y = \sigma_0^{b_n}(z)$, with z a Sturmian sequence. Then y contains the words $10^{b_n}1$ and $10^{b_n}0$.*

Proof. If $y = \sigma_0^{b_n}(z)$, then y is of type 0 and z is of type 1. By iteration, it is straightforward that $\sigma_0^{b_n}(0) = 0$ and $\sigma_0^{b_n}(1) = 10^{b_n}$. Since z is of type 1, it contains the words 10 and 11. Therefore, y contains the words:

$$10^{b_n}10^{b_n} \quad \text{and} \quad 10^{b_n}0.$$

\square

Proof of Proposition 12: Let x be a Sturmian sequence, with α the frequency of 1, and assume that α does not have bounded type. Then its multiplicative coding $(b_n)_{n \in \mathbb{N}}$ is an unbounded sequence.

Let $y = \Phi^{b_0+b_1+\dots+b_{n-1}}(x)$, and $z = \Phi^{b_n}(y)$. Without loss of generality, we assume that z is of type 1, so that

$$y = \sigma_0^{b_n}(z).$$

Then, using previous lemma, y contains the words $10^{b_n}1$ and $10^{b_n}0$. Therefore, for all $1 \leq k \leq [b_n]/2 - 1$, y contains the words $00^{2k}0$ and $00^{2k}1$.

Applying the substitution $\sigma := \sigma_0^{b_1} \circ \dots \circ \sigma_0^{b_{n-1}}$, the sequence x contains the words

$$\sigma(0)\sigma(0)^{2k}\sigma(0) \quad \text{and} \quad \sigma(0)\sigma(0)^{2k}\sigma(1).$$

Let a be the last letter of $\sigma(0)$ and $\sigma(1)$ (it is the same: 0 if $b_0 \neq 0$, 1 otherwise). Since $\sigma(1)$ starts by 1 and $\sigma(0)$ starts by 0, x contains the words

$$(11) \quad a\sigma(0)^k\sigma(0)^k0 \quad \text{and} \quad a\sigma(0)^k\sigma(0)^k1.$$

Then, let $X_k = [a\sigma(0)^k \cdot \sigma(0)^k]$, where the dot separates the indices $i \leq 0$ and $i > 0$. It is an element of $\mathcal{P}_{k|\sigma(0)|}$, where $|\sigma(0)|$ is the length of $\sigma(0)$. Let v_k be the associated vertex. Note that v_{k+1} is a child of v_k for all k . From Equation (11), the vertices v_k have two distinct children (in the non-reduced tree of words), therefore, they are elements of the reduced tree, and their weight is $(k|\sigma(0)|)^{-1}$. In particular, this construction shows that there are two vertices u, v (namely v_1 and $v_{[b_n]/2-1}$), such that the quotient of their weights is $[b_n/2] - 1$, and their distance in the tree is $[b_n/2] - 1$.

This construction can be done for all n . If $(b_n)_{n \in \mathbb{N}}$ is unbounded, this shows that the weights cannot satisfy the geometric decay condition of Proposition 4, and Ξ is not embeddable. \square

The results presented here (namely proposition 11 and 12) provide a proof of Theorem 2. It is actually possible to give a more complete version of it.

Theorem 11. *A Sturmian subshift $\Xi(\alpha)$ is f -embeddable if and only if the irrational number α associated with it has bounded type. In particular*

- (i) *if α is a quadratic irrational, then $\Xi(\alpha)$ is f -embeddable;*
- (ii) *for almost every $\alpha \in (0, 1)$, the subshift $\Xi(\alpha)$ is not f -embeddable;*
- (iii) *the boundary of $\Xi(e)$ is not embeddable for $e = 2.71828\dots$*

Proof. The first part of this result is Theorem 2 which was proved above. Point (i) is a consequence of the fact that quadratic irrational have an eventually periodic (hence bounded) continued fraction expansion. Point (ii) is a consequence of the theorem of Khintchine on continued fractions, and point (iii) is a consequence of the explicit formula for the continued fraction decomposition of e . \square

6. HAUSDORFF DIMENSION

The Assouad dimension provides a good measure of embeddability of ultrametric Cantor spaces in \mathbb{R}^n . In [19], it was proved that for *self-similar* Cantor sets, the smaller dimension of a space \mathbb{R}^n in which an ultrametric Cantor (C, d) set can be embedded is equal to $\lfloor \dim_{\text{H}}(C, d) \rfloor + 1$, where \dim_{H} is the Hausdorff dimension. When the space is not self-similar however, there is no reason that the Hausdorff dimension has anything to do with the possibility to embed the space. For example [25] mentions an example of a countable metric space (hence of Hausdorff dimension 0) which cannot be bi-Lipschitz embedded in any finite-dimensional Euclidean space. The goal of this section is to revisit the definition and properties of Hausdorff dimension in terms of the Michon representation for ultrametric Cantor sets. Proposition 14 gives an adapted formula for the computation of the Hausdorff dimension, which will be used in next section.

6.1. Hausdorff Dimension of the boundary of a tree. The definition of the Hausdorff dimension [14] starts with the following construction: given an open cover \mathcal{U} of $C = \partial\mathcal{T}$ and given $s \in [0, \infty)$, let $\mathcal{H}^s(\mathcal{U})$ be defined by

$$(12) \quad \mathcal{H}^s(\mathcal{U}) = \sum_{U \in \mathcal{U}} \text{diam}(U)^s.$$

In addition, let $\text{diam}(\mathcal{U}) = \sup_{U \in \mathcal{U}} \text{diam}(U)$. Then for $0 < \delta < 1$ let $\mathcal{H}_\delta^s(C)$ be defined by

$$(13) \quad \mathcal{H}_\delta^s(C) = \inf_{\text{diam}(\mathcal{U}) < \delta} \mathcal{H}^s(\mathcal{U})$$

It follows that $\delta' \leq \delta \Rightarrow \mathcal{H}_\delta^s(C) \leq \mathcal{H}_{\delta'}^s(C)$. In addition, if $\sigma > 0$ then $\mathcal{H}_\delta^{s+\sigma}(C) \leq \delta^\sigma \mathcal{H}_\delta^s(C) \leq \mathcal{H}_\delta^s(C)$. Consequently,

- (i) the limit $\lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(C) = \mathcal{H}^s(C)$ exists in $[0, +\infty) \cup \{+\infty\}$,
- (ii) there is a unique s_0 such that if $s > s_0$ then $\mathcal{H}^s(C) = 0$ whereas for $s < s_0$, $\mathcal{H}^s(C) = +\infty$.

This unique value s_0 is precisely the Hausdorff dimension $s_0 = \dim_{\text{H}}(C, d)$. In order to give a more tractable formula, the following definition will be needed

Definition 12. *A finite subtree Γ of \mathcal{T} is a tree graph $\Gamma = (\mathcal{V}_\Gamma, \mathcal{E}_\Gamma, \bullet)$ where*

- (i) $\mathcal{V}_\Gamma \subset \mathcal{V}$ is finite and contains the root \bullet ;
- (ii) for every $v \in \mathcal{V}_\Gamma$ every ancestor of v belong to \mathcal{V}_Γ ;
- (iii) the “parent” relation is induced by the one on \mathcal{T} (equivalently, the edges of Γ are the edges of \mathcal{T} between vertices of Γ).

$\mathcal{L}\Gamma$ will denote the set of leaves (vertices of maximal depth). The finite subtree Γ is called full if for any vertex $v \in \Gamma \setminus \mathcal{L}\Gamma$, each child of v is a vertex of Γ . It is equivalent to the statement $\partial\mathcal{T} = \bigcup_{v \in \mathcal{L}\Gamma} [v]$.

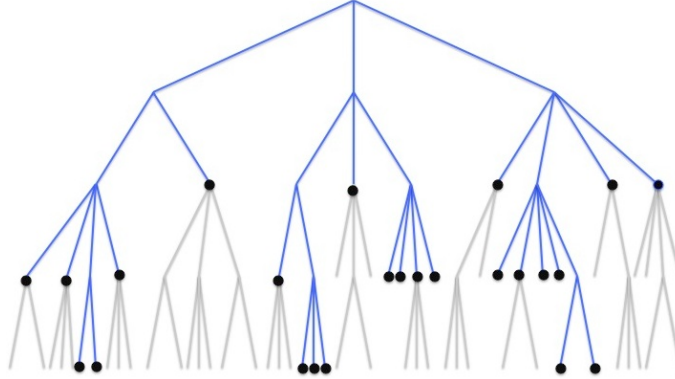


FIGURE 1. A full finite subtree and its extremal vertices

The following theorem allows for simpler computation of the Hausdorff dimension.

Proposition 14. *Let \mathcal{T} be a reduced Michon tree with weight κ . For any $\delta > 0$ let $\mathfrak{G}_\delta(\mathcal{T})$ be the set of all full finite subtrees of \mathcal{T} such that $\max_{v \in \mathcal{L}\Gamma} \kappa(v) < \delta$. Then*

$$(14) \quad \mathcal{H}_\delta^s(\partial\mathcal{T}) = \inf_{\Gamma \in \mathfrak{G}_\delta(\mathcal{T})} \sum_{v \in \mathcal{L}\Gamma} \kappa(v)^s.$$

The following theorem was known, even for more general metric spaces than ultrametric ones. With our notations, the proof becomes very simple and is included.

Corollary 1 (Assouad [4], Prop. 2-(j)). *Let (X, d) be an ultrametric Cantor set. Then*

$$\dim_{\text{H}}(X, d) \leq \dim_{\text{A}}(X, d).$$

Conversely, for any reduced tree \mathcal{T} , there exists a weight which turns $\partial\mathcal{T}$ into an ultrametric Cantor set of Hausdorff dimension 1. It is the Kraft weight defined inductively by $\kappa(\bullet) = 1$ and for all w , if one notes v the parent of w and assume that v has k children, then $\kappa(w) = \kappa(v)/k$. It leads to the following corollary, the proof of which is left to the reader,

Corollary 2. *Let \mathcal{T} be a reduced Michon tree with its Kraft weight K . The Hausdorff dimension of $(\partial\mathcal{T}, d_K)$ is exactly one. In particular $(\partial\mathcal{T}, d_K)$ is f -embeddable if and only if \mathcal{T} has a bounded number of children per vertex.*

The proofs of the first two results will be the content of the following section.

6.2. Proof of Proposition 14. Remark that the right-hand side of formula (14) is an infimum of the quantity $\sum_{u \in \mathcal{U}} \text{diam}(U)^s$ taken over the special partitions of the form $\mathcal{U} = \{[v] ; v \in \mathcal{L}\Gamma\}$. Therefore,

$$(15) \quad \mathcal{H}_\delta^s(\partial\mathcal{T}) \leq \inf_{\Gamma \in \mathfrak{G}_\delta(\mathcal{T})} \sum_{v \in \mathcal{L}\Gamma} \kappa(v)^s.$$

Conversely, let \mathcal{U} be a cover of $\partial\mathcal{T}$ by open sets of diameter less than δ . By definition of the distance on $\partial\mathcal{T}$, for each $U \in \mathcal{U}$, there is a vertex v_U such that $U \subset [v_U]$ and $\kappa(v_U) = \text{diam}(U)$.

The family $(v_U)_{U \in \mathcal{U}}$ form a covering of $\partial\mathcal{T}$ by clopen sets of diameter less than δ . By compactness, there is a finite sub-cover $[v_1], \dots, [v_m]$. Then, for each pair $v_i \neq v_j$, exactly one of the following holds: (a) v_i is an ancestor of v_j (in which case $[v_j] \subset [v_i]$); (b) or the converse holds; (c) or neither is an ancestor of the other (in which case $[v_i] \cap [v_j] = \emptyset$). By removing all clopen sets in this covering which are included in others, it is possible to get a finite *partition* by clopen sets $([v'_1], \dots, [v'_n])$. In the process, terms have been removed. Therefore

$$\sum_{U \in \mathcal{U}} \text{diam}(U)^s = \sum_{U \in \mathcal{U}} \kappa(v_U)^s \geq \sum_{i=1}^n \kappa(v'_i)^s.$$

It is straightforward to see that the subtree consisting of vertices v'_1, \dots, v'_n and all their ancestors is a finite full subtree of \mathcal{T} . Therefore, for each covering of $\partial\mathcal{T}$ by open sets, there is a full finite subtree Γ such that

$$\sum_{U \in \mathcal{U}} \text{diam}(U)^s \geq \sum_{v \in \mathcal{L}\Gamma} \kappa(v)^s.$$

It proves that inequality (15) is in fact an equality, which achieves the proof.

6.3. Proof of Corollary 1. Let (X, d) be an ultrametric Cantor set represented by its reduced Michon tree \mathcal{T} with weight κ . Without loss of generality, the weight of the root $\kappa(\delta)$ can be taken to be 1. If (X, d) is s -subhomogeneous, then, for any $\varepsilon > 0$, $\mathcal{H}^{s+\varepsilon}(X) = 0$. For indeed, using Proposition 3, there is C such that for all δ , the following holds (where \bullet is the root of \mathcal{T}):

$$\#(\mathcal{L}(\mathcal{T}, \bullet, \delta)) \leq C\delta^{-s}.$$

It ought to be remarked that $\mathcal{L}(\mathcal{T}, \bullet, \delta)$ is nothing but the leaves of the full finite subtree Γ consisting precisely of these leaves and all of their ancestors up to the root. By definition, these leaves define clopen sets $[v]$ of diameter less than δ . Therefore by Proposition 14,

$$\mathcal{H}_\delta^{s+\varepsilon}(X) \leq \sum_{v \in \mathcal{L}\Gamma} \kappa(v)^{s+\varepsilon}.$$

Now, the number of terms in the sum is bounded above by $C\delta^{-s}$. Therefore $\mathcal{H}_\delta^{s+\varepsilon}$ is bounded above by $C\delta^\varepsilon$, which tends to 0 as δ tends to 0.

7. RANDOM TREES

Random trees are an ubiquitous objects in modern mathematics. It is therefore natural to investigate the boundary of such trees, at they are good candidates to provide examples of Cantor sets with typical properties. Originally, Sir Francis Galton and the reverend Watson introduced these trees to investigate the probability of disappearance of family names among the British aristocrats in the nineteenth century. Nowadays, Galton-Watson processes appear in situations as varied as, for example, the description of the nuclear chain reactions or of electron emission in a photomultiplier tube (see [16]).

A Galton-Watson branching process, as it is called today [16, 6] defines naturally a tree. Starting with a root \bullet , its number of children ξ_\bullet is a random variable with integer value and probability distribution p . Each offspring is represented by a vertex v and each vertex has a random variable ξ_v attached to it, describing the number of its children. All ξ_v are independent and identically distributed, with distribution p . Going inductively in the same way, leads to a tree. It is finite with positive probability if $p_0 := p(\{0\}) > 0$ and is almost surely infinite otherwise. In order to fit in the setting of this paper, the resulting tree ought to be infinite and reduced namely

each vertex should have at least two children. In order to ensure it, it will be assumed that $p_0 = p_1 = 0$. The resulting branching process is called a *reduced random tree*.

It is worth noting that, using the deterministic weight $\kappa(v) = 2^{-|v|}$, the dimension of the boundary of trees obtained by a Galton-Watson process was computed by Hawkes [17] to be given by $\log(m)/\log(2)$, where $m > 1$ is the average number of children per vertex. The Hausdorff dimension associated with this metric can be used to define an “average branching number”-even for more general trees- which turns out to be the crucial parameter when studying the behavior of random walks on trees [26]. The approach taken in the present paper is however different, since the trees are endowed with random weights rather than a deterministic one. To the authors best knowledge, this approach is new.

7.1. Reduced Random Trees: a review. Let $p = (p_n)_{n \geq 2}$ be the probability supported by $[2, \infty) \subset \mathbb{N}$ describing the number of children of each vertex. Let Z_n be the number of descendant at the generation n . Following Watson’s idea, it is convenient to introduce the generating function

$$P_n(x) = \sum_{l=2}^{\infty} \text{Prob}\{Z_n = l\} x^l$$

The construction of the random tree implies the following formula, where \mathcal{V}_n denotes the set of vertices at generation n ,

$$(16) \quad Z_{n+1} = \sum_{v \in \mathcal{V}_n} \xi_v.$$

Consequently the conditional probabilities are given by

$$\text{Prob}\{Z_{n+1} = l \mid Z_n = k\} = \sum_{j_1 + \dots + j_k = l} p_{j_1} \cdots p_{j_k}.$$

In particular, it shows that $(Z_n)_{n \in \mathbb{N}_*}$ defines a Markov chain. Moreover

$$(17) \quad P_{n+1}(x) = P_n(P(x)), \quad P(x) = \sum_{n=2}^{\infty} p_n x^n.$$

Since p is a probability, the series defining P converges for $0 \leq x \leq 1$. In addition, $P(1) = 1$, $m = P'(1) = \mathbb{E}(\xi)$ represents the average number of offsprings, if it exists. With the present restrictions, it follows that $m \geq 2$. The function $x \in [0, 1] \rightarrow P(x) \in [0, 1]$ is positive, increasing and convex, more generally all its derivatives, when they exist, are positive, meaning that P is *completely monotone*. From the recursion relation (17) it follows immediately that

$$P_n(x) = \underbrace{P \circ P \circ \dots \circ P}_n(x)$$

In particular, since $P(1) = 1$, it follows that $\mathbb{E}(Z_n) = P'_n(1) = m^n$ for all n . Hence the number of descendants at generation n grows exponentially fast with the generation in the average.

Let now \mathcal{F}_n be the sigma algebra generated by the variables ξ_v for v vertices of the generations $k \leq n$. A classical remark made by Doob [10], is that, thanks to the equation (16) defining the process, the family $(Z_n)_{n \in \mathbb{N}_*}$ satisfies

$$\mathbb{E}(Z_{n+1}|\mathcal{F}_n) = mZ_n.$$

Namely $W_n = Z_n/m^n$ is a *martingale*. As shown earlier by Doob [9] this implies

Theorem 12 (See [16, 6]). *Let $p = (p_n)_{n \geq 2}$ be the probability distribution for the number ξ of offsprings such that the average $\mathbb{E}(\xi) = m$ and the variance $\text{Var}(\xi) = \sigma^2$ are finite. Then the sequence $W_n = Z_n/m^n$ of random variables is a martingale with respect to the increasing sequence \mathcal{F}_n of σ -algebras. In particular it converges almost surely to a random variable W such that*

$$\mathbb{E}(W) = 1, \quad \text{Var}(W) = \frac{\sigma^2}{m^2 - m}.$$

7.2. Proof of the Proposition 1. If the probability p has an infinite support, given any integer $M \geq 2$, the probability P_M that a given vertex has more than M children is non zero. The construction of the Random Reduced Tree, can be seen by associating inductively with any vertex of generation n a string (b_1, b_2, \dots, b_n) of integers so that $1 \leq b_n \leq \xi_{b_1, \dots, b_{n-1}}$. In particular, since $\xi_v \geq 2$ for all v 's almost surely the subset $\mathcal{W} \subset \mathcal{V}$ made of vertices for which all b_j 's belong to $\{1, 2\}$ is non empty and gives an infinite binary subtree. Since the random variables $\{\xi_v; v \in \mathcal{W}\}$ are *i.i.d.*, it follows that , given $M \geq 2$

$$\text{Prob}\{\xi_v \leq M; \forall v \in A\} = (1 - P_M)^{\#A}.$$

In particular the probability that all vertices in \mathcal{W} have $\xi_v \leq M$ vanishes. \square

7.3. Random Weight. In order that the boundary of the rooted random tree $\mathcal{T} = (\mathcal{V}, \mathcal{E}, \bullet)$ previously built becomes an ultrametric Cantor set, it is necessary to put a weight on each vertex. The previous construction suggests that the weight itself be random and Markovian as well. In order to do so, the following model of random weight is proposed: let $(\lambda_v)_{v \in \mathcal{V}}$ be a family of *i.i.d* in $[0, 1]$ with common distribution ρ ; then the weight $\kappa(v)$ will be given by

$$\kappa(v) = \lambda_v \kappa(u), \quad \text{if } u = \text{parent of } v, \quad \kappa(\bullet) = 1.$$

In order that this defines a good weight, it is required that $\lambda_v \neq 0$ with probability one, namely $\rho\{0\} = 0$. Moreover, in order that the weight converges to zero along any infinite paths with probability one, it will be required that $\rho\{1\} < 1$. It will be convenient to use the following generating function (Mellin transform)

$$h(s) = \int_0^1 \lambda^s \rho(d\lambda).$$

It is easy to see that h is decreasing, logarithmically convex, namely $h(\eta s_1 + (1 - \eta)s_0) \leq h(s_1)^\eta h(s_0)^{1-\eta}$ for $0 < \eta < 1$, and $\lim_{s \rightarrow \infty} h(s) = \rho\{1\}$.

7.4. Proof of Theorem 3. Thanks to Proposition 14, the Hausdorff dimension of the tree can be computed from the following random variables

$$(18) \quad \mathcal{H}^s(\Gamma) = \sum_{v \in \mathcal{L}\Gamma} \kappa(v)^s, \quad \mathcal{H}_n^s = \sum_{v \in \mathcal{V}_n} \kappa(v)^s.$$

where Γ is a full finite subtree of the random \mathcal{T} . In the following \mathfrak{G} will denote the set of finite full subtrees of \mathcal{T} . This set is ordered by the inclusion of the vertex sets. In addition, giving

$\Gamma, \Gamma' \in \mathfrak{G}$, then they admit a least upper bound $\Gamma \vee \Gamma'$ and a greatest lower bound $\Gamma \wedge \Gamma'$, as can be checked immediately. This suggests to define a new family of σ -algebras: let $\widehat{\mathcal{F}}_\Gamma$ be the σ -algebra generated by the ξ_u 's and the λ_v 's where $v \in \mathcal{V}_\Gamma$ and $u \in \mathcal{V}_\Gamma \setminus \mathcal{L}\Gamma$. In particular, if Γ is the full tree associated with generation n , then $\widehat{\mathcal{F}}_n$ will denote the σ -algebra generated by the ξ_u 's with u vertex of generation $k \leq n-1$ and by the λ_v 's with v vertex of generation $k \leq n$. With these notations, the following result holds

Proposition 15. *With the assumptions made previously on the construction of the random rooted weighted tree $(\mathcal{T}, \mathcal{V}, \bullet, \kappa)$ the following results hold*

(i) *For all $s \geq 0$, the family $Y_n(s) = \mathcal{H}_n^s / m^n h(s)^n$ is a martingale with respect to the family $(\widehat{\mathcal{F}}_n)_{n \in \mathbb{N}}$ of σ -algebras. In particular it converges almost surely to a positive random variable $Y(s)$ such that $\mathbb{E}(Y(s)) = 1$.*

(ii) *There is $t_m > s_m$ defined as the unique solution of $h(2s) = mh(s)^2$, such that*
 (a) *if $s < t_m$ then $Y_n(s)$ converges almost surely to a constant,*
 (b) *if $s = t_m$, then,*

$$\text{Var}(Y(t_m)) = 1 - \frac{1}{m} + \frac{\sigma^2}{m^2},$$

(c) *if $s > t_m$ the random variable $Y(s)$ has not a finite second moment.*

(iii) *If $\rho\{1\} < m^{-1}$ and if s_m is the unique solution of $mh(s) = 1$, then the sequence of random variables $Y(\Gamma) = \mathcal{H}^{s_m}(\Gamma)$ defines a martingale with respect to the family $(\widehat{\mathcal{F}}_\Gamma)_{\Gamma \in \mathfrak{G}}$ of σ -algebras. In particular it converges almost surely to 1.*

Proof. (i) From the definition of \mathcal{H}_n^s in eq. (18), it follows that

$$\mathbb{E}(\mathcal{H}_1^s) = \mathbb{E}\left(\sum_{v \in \text{Ch}(\bullet)} \lambda^s\right) = h(s)\mathbb{E}(\xi_\bullet) = mh(s).$$

Moreover,

$$\begin{aligned} \mathbb{E}(\mathcal{H}_{n+1}^s | \widehat{\mathcal{F}}_n) &= \mathbb{E}\left(\sum_{u \in \mathcal{V}_n} \kappa(u)^s \sum_{v \in \text{Ch}(u)} \lambda_v^s \middle| \widehat{\mathcal{F}}_n\right) \\ &= h(s)\mathbb{E}\left(\sum_{u \in \mathcal{V}_n} \kappa(u)^s \xi_u \middle| \widehat{\mathcal{F}}_n\right) = mh(s) \sum_{u \in \mathcal{V}_n} \kappa(u)^s \\ &= mh(s)\mathcal{H}_n^s. \end{aligned}$$

This calculation shows that $\mathbb{E}(Y_{n+1}(s) | \widehat{\mathcal{F}}_n) = Y_n(s)$. In particular it is a martingale *w.r.t.* the family $\widehat{\mathcal{F}}_n$ of σ -algebras. Since $Y_1(s) = \mathcal{H}^s / mh(s)$ it follows that $\mathbb{E}(Y_1(s)) = 1$. Therefore $\mathbb{E}(Y_n(s)) = 1$ for all n 's. The convergence of this family is the main result of the martingale theory [10].

(ii) The calculation of the variance will be done through the second moment of \mathcal{H}_n^s . By construction

$$\mathbb{E} \left((\mathcal{H}_{n+1}^s)^2 \mid \widehat{\mathcal{F}}_n \right) = \mathbb{E} \left(\sum_{u, u' \in \mathcal{V}_n} \kappa(u)^s \kappa(u')^s \sum_{v \in \text{Ch}(u)} \sum_{v' \in \text{Ch}(u')} \lambda_v^s \lambda_{v'}^s \mid \widehat{\mathcal{F}}_n \right).$$

Let the terms with $u \neq u'$ be considered first. Then, since the λ_v 's are independent for different v 's, it follows that

$$\begin{aligned} \mathbb{E} \left(\kappa(u)^s \kappa(u')^s \sum_{v \in \text{Ch}(u)} \sum_{v' \in \text{Ch}(u')} \lambda_v^s \lambda_{v'}^s \mid \widehat{\mathcal{F}}_n \right) &= h(s)^2 \mathbb{E} \left(\kappa(u)^s \kappa(u')^s \xi_u \xi_{u'} \mid \widehat{\mathcal{F}}_n \right) \\ &= h(s)^2 m^2 \kappa(u)^s \kappa(u')^s. \end{aligned}$$

If now $u = u'$, this gives two terms: the first one are terms for which $v \neq v'$ and the other ones are for $v = v'$. The same type of calculation leads to

$$\begin{aligned} \mathbb{E} \left(\kappa(u)^{2s} \sum_{v \neq v' \in \text{Ch}(u)} \lambda_v^s \lambda_{v'}^s \mid \widehat{\mathcal{F}}_n \right) &= h(s)^2 (m^2 + \sigma^2 - m) \kappa(u)^{2s}, \\ \mathbb{E} \left(\kappa(u)^{2s} \sum_{v \in \text{Ch}(u)} \lambda_v^{2s} \mid \widehat{\mathcal{F}}_n \right) &= h(2s) m \kappa(u)^{2s}. \end{aligned}$$

Grouping these results, leads to

$$\mathbb{E} \left((\mathcal{H}_{n+1}^s)^2 \mid \widehat{\mathcal{F}}_n \right) = h(s)^2 m^2 (\mathcal{H}_n^s)^2 + \{m (h(2s) - h(s)^2) + \sigma^2 h(s)^2\} \mathcal{H}_n^{2s}.$$

Averaging on both sides gives

$$\text{Var}(\mathcal{H}_{n+1}^s) = \{m (h(2s) - h(s)^2) + \sigma^2 h(s)^2\} m^n h(2s)^n.$$

It is worth noticing that, thanks to the definition of h , the Cauchy-Schwarz inequality gives $h(s)^2 < h(2s)$. This inequality is actually strict because $\rho\{1\} \neq 1$, so that λ is not almost surely equal to one. Therefore

$$\text{Var}(Y_n(s)) = \left(\frac{1}{m} \frac{h(2s)}{h(s)^2} \right)^n \left(1 - \left(1 - \frac{\sigma^2}{m} \right) \frac{h(s)^2}{h(2s)} \right).$$

It follows that, if $s < s_m$, then $h(s)m > 1$. In addition an elementary calculation shows that the map $g(s) = h(2s)/h(s)^2$ is monotone increasing, that $g(0) = 1$ and $\lim_{s \rightarrow \infty} g(s) = \rho\{1\}^{-1}$. Therefore, there is a unique $t_m > 0$ such that $m = g(t_m)$. Using the definition of s_m , it is easy to show that $s_m < t_m$. Hence

(a) if $s < t_m$, $\lim_{n \rightarrow \infty} \text{Var}(Y_n(s)) = 0$, implying that Y_n converges almost surely to a constant; this constant can only be the common average, namely $\lim_{n \rightarrow \infty} Y_n(s) = 1$,

(b) if $s = t_m$, then the variance converges to a finite value

$$s = t_m \Rightarrow \lim_{n \rightarrow \infty} \text{Var}(Y_n(t_m)) = \left(1 + \frac{\sigma^2}{m^2} - \frac{1}{m} \right),$$

(c) if $s > t_m$, then the limiting random variable $Y(s)$ does not have a finite second moment.

(iii) Let now $\Gamma' \subset \Gamma$ be two full finite subtrees of \mathcal{T} . Then there is a decreasing sequence of full finite subtrees such that $\Gamma' \subset \Gamma_j \subset \cdots \Gamma_1 \subset \Gamma_0 = \Gamma$, and such that Γ_{i+1} is obtained from Γ_i by the following procedure: each vertex $v \in \mathcal{L}\Gamma_i$ which is not in $\mathcal{L}\Gamma'$ is removed and replaced by its parent. It is clear that, if Γ_i is full, so is Γ_{i+1} . This leads to

$$\mathbb{E} \left(\mathcal{H}^s(\Gamma) \mid \widehat{\mathcal{F}}_{\Gamma_1} \right) = \mathbb{E} \left(\sum_{u \in \mathcal{L}\Gamma_1; \text{Ch}(u) \cap \mathcal{L}\Gamma' = \emptyset} \kappa(u)^s \sum_{v \in \text{Ch}(u)} \lambda_v^s \mid \widehat{\mathcal{F}}_{\Gamma_1} \right) + \sum_{u \in \mathcal{L}\Gamma \cap \mathcal{L}\Gamma'} \kappa(u)^s$$

Thanks to the definition of $\widehat{\mathcal{F}}_{\Gamma_1}$ the *r.h.s.* becomes

$$\mathbb{E} \left(\mathcal{H}^s(\Gamma) \mid \widehat{\mathcal{F}}_{\Gamma_1} \right) = mh(s) \sum_{u \in \mathcal{L}\Gamma_1; \text{Ch}(u) \cap \mathcal{L}\Gamma' = \emptyset} \kappa(u)^s + \sum_{u \in \mathcal{L}\Gamma \cap \mathcal{L}\Gamma'} \kappa(u)^s.$$

In particular, if $s = s_m$, namely if $mh(s) = 1$, this gives $Y(\Gamma) = \mathcal{H}^{s_m}(\Gamma)$ so that

$$\mathbb{E} \left(Y(\Gamma) \mid \widehat{\mathcal{F}}_{\Gamma_1} \right) = Y(\Gamma_1).$$

Proceeding inductively along the chain of Γ_i 's, this gives

$$\mathbb{E} \left(Y(\Gamma) \mid \widehat{\mathcal{F}}_{\Gamma'} \right) = Y(\Gamma').$$

Therefore the family $(Y(\Gamma))_{\Gamma \in \mathfrak{G}}$ is also a martingale *w.r.t.* the $\widehat{\mathcal{F}}_{\Gamma}$'s. The martingale theorem then shows that it converges almost surely. Since the full tree with boundary \mathcal{V}_n is a member of this family and since it has been shown that the variance converges to zero (because $s_m < t_m$), the family converges to a constant almost surely. \square

Proposition 16. *Under the hypothesis of Proposition 15, the Hausdorff dimension of $(\partial\mathcal{T}, d_\kappa)$ is almost surely equal to s_m .*

Proof. Thanks to Proposition 15, $\mathcal{H}^{s_m}(\Gamma)$ converges almost surely to 1. It follows that

$$\mathcal{H}_\delta^{s_m} = \inf_{\Gamma \in \mathfrak{G}_\delta} \mathcal{H}^{s_m}(\Gamma), \quad \Rightarrow \quad \lim_{\delta \downarrow 0} \mathcal{H}_\delta^{s_m} = 1.$$

Consequently, $\mathcal{H}_\delta^s \rightarrow \infty$ for $s < s_m$ and $\mathcal{H}_\delta^s \rightarrow 0$ for $s > s_m$. Hence $\dim_{\text{H}}(\partial\mathcal{T}, d_\kappa) = s_m$. \square

Proposition 17. *Under the hypothesis of Proposition 15, the Hausdorff measure of $(\partial\mathcal{T}, d_\kappa)$ at the dimension $s = s_m$ exists almost surely and is a random probability.*

Proof. In order to prove it, it is sufficient to consider the basis of clopen sets of the form $[u]$ for $u \in \mathcal{V}$. It boils down to consider

$$\mathcal{H}^{s_m}(\Gamma; u) = \sum_{v \in \mathcal{L}\Gamma; v \preceq u} \kappa(v)^{s_m}.$$

a calculation similar to the one made in the proof of Proposition 15, shows that the family of $\{\mathcal{H}^{s_m}(\Gamma; u); \Gamma \in \mathfrak{G}, u \in \mathcal{V}_\Gamma\}$ is also a martingale satisfying

$$\mathbb{E} \left(\mathcal{H}^{s_m}(\Gamma; u) \mid \widehat{\mathcal{F}}_{\Gamma_0} \right) = \kappa(u)^{s_m},$$

for all full finite subtree Γ_0 with $u \in \mathcal{L}\Gamma_0$ and $\Gamma \supset \Gamma_0$. Therefore, the Martingale Theorem implies that $\mu([u]) = \lim_{\Gamma} \mathcal{H}^{s_m}(\Gamma; u)$ exists and that it is a random variable. Since the set of

vertices is countable, the set of probability zero on which the convergence does not hold can be chosen independently on $u \in \mathcal{V}$. By construction

$$\sum_{u \in \mathcal{L}\Gamma_0} \mathcal{H}^{s_m}(\Gamma; u) = \mathcal{H}^{s_m}(\Gamma),$$

showing that, after taking the limit, $\sum_{u \in \mathcal{L}\Gamma_0} \mu([u]) = 1$. In addition, whenever Γ is a full finite subtree, $\{[v] ; v \in \mathcal{L}\Gamma_0\}$ forms a partition by clopen sets. Moreover, clopen sets of the form $[v]$ (v a vertex of the tree) generate the σ -algebra of Borel sets. Hence μ defines a probability measure on $\partial\mathcal{T}$. \square

REFERENCES

- [1] J. Aliste, D. Coronel, “Tower systems for linearly repetitive Delone sets”, *Ergod. Th. & Dynam. Sys.*, **31**, (2011), 1595-1618.
- [2] P. Arnoux, “Sturmian Sequences”, Chap. 6 in N. Pytheas Fogg, *Substitutions in dynamics, arithmetics, and combinatorics*, Lecture Notes in Mathematics, Springer, (2002).
- [3] P. Assouad, “Espaces métriques, plongements, facteurs.”, PhD thesis, *Publications Mathématiques d’Orsay*, (1977) No. 223-7769.
- [4] P. Assouad, “Étude d’une dimension métrique liée à la possibilité de plongements dans \mathbb{R}^n ”, *C. R. Acad. Sci. Paris Sér. A-B*, **288**, (1979), no. 15, A731-A734.
- [5] P. Assouad, “Plongements lipschitziens dans \mathbb{R}^n ”, *Bull. Soc. Math. France*, **111**, (1983), no. 4, 429-448.
- [6] K. B. Athreya, P.E. Ney, *Branching processes*, Die Grundlehren der mathematischen Wissenschaften, Band 196. Springer-Verlag, New York-Heidelberg, (1972). Reprinted by Dover Publications, Inc., Mineola, NY, (2004).
- [7] J. Bellissard, R. Benedetti, J.-M. Gambaudo, “Spaces of tilings, finite telescopic approximations and gap-labeling”, *Commun. Math. Phys.*, **261**, (2006), 1-41.
- [8] L. Brouwer, “On the structure of perfect sets of points”, *Proc. Akad. Amsterdam*, **12**, (1910), 785-794.
- [9] J. L. Doob, “Regularity properties of certain families of chance variables”, *Trans. Amer. Math. Soc.*, **47**, (1940), 455-486.
- [10] J. L. Doob, *Stochastic processes*, John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London, (1953).
- [11] F. Durand, “Linearly recurrent subshifts have a finite number of non-periodic subshift factors”, *Ergod. Th. & Dynam. Sys.* **20**, (2000), 1061-1078.
- [12] F. Durand, “Corrigendum and addendum to ‘Linearly recurrent subshifts have a finite number of non-periodic factors’”, *Ergod. Th. & Dynam. Sys.* **23**, (2003), 663-669.
- [13] F. Durand, B. Host, C. Skau, “Substitutions, Bratteli diagrams and dimension groups”, *Ergod. Th. & Dynam. Sys.* **19**, (1999), 953-993.
- [14] K. Falconer, *Fractal Geometry, Mathematical foundations and applications*, John Wiley & Sons, Ltd., Chichester, (1990).
- [15] M. Fréchet, “Les dimensions d’un ensemble abstrait”, *Mathematische Annalen*, **68**, (1910), 145-168.
- [16] T. E. Harris, *The theory of branching processes*, Die Grundlehren der Mathematischen Wissenschaften, Bd. 119 Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, N.J. (1963). *The theory of branching processes* (Corrected reprint of the 1963 original book), Dover Phoenix Editions, Dover Publications, Inc., Mineola, NY, (2002).
- [17] J. Hawkes, “Trees generated by a simple branching process”, *J. London Math. Soc. (2)*, **24** (1981), no. 2, 373-384.
- [18] F. Hippert & D. Gratias Eds., *Lectures on Quasicrystals*, Editions de Physique, Les Ulis, (1994).
- [19] A. Julien, J. Savinien, “Embeddings of self-similar ultrametric Cantor sets”, *Topology and Applications*, **158**, (2011), 2148-2157.
- [20] A. Julien, J. Savinien, “Transverse Laplacians for substitution tilings”, *Comm. Math. Phys.*, **301**, (2011), 285-318.
- [21] D. G. Kendall, “Branching processes since 1873”, *J. London Math. Soc.*, **41**, (1966), 385-406.
- [22] A. Khintchine, *Continued Fractions*, New York: Dover, (1997).
- [23] P. Lévy, “Sur le développement en fraction continue d’un nombre choisi au hasard”, (French) *Compositio Math.*, **3**, (1936), 286-303.
- [24] K. Luosto, “Ultrametric spaces bi-Lipschitz embeddable in \mathbb{R}^n ”, *Fund. Math.*, **150**, (1996), no. 1, 25-42.
- [25] J. Luukkainen, H. Movahedi-Lankarani, “Minimal bi-Lipschitz embedding dimension of ultrametric spaces”, *Fund. Math.*, **144** (1994), no. 2, 181-193.
- [26] R. Lyons, “Random walks, capacity and percolation on trees”, *Ann. Probab.*, **20**, (1992), no. 4, 2043-2088.
- [27] G. Michon, “Les Cantors réguliers”, *C. R. Acad. Sci. Paris Sér. I Math.*, **300**, (1985), 673-675.
- [28] M. Morse, G. A. Hedlund, “Symbolic Dynamics II. Sturmian trajectories”, *American Journal of Mathematics*, **62**, (1940), 1-42.
- [29] H. Movahedi-Lankarani, “On the inverse of Mañé’s projection”, *Proc. of the A.M.S.*, **116**, (1992), 555-560
- [30] J. Pearson, J. Bellissard, “Noncommutative Riemannian Geometry and Diffusion on Ultrametric Cantor Sets”, *Journal of Noncommutative Geometry*, **3**, (2009), 447-480.

- [31] I. J. Schoenberg, “On certain metric spaces arising from euclidean spaces by a change of metric and their imbedding in Hilbert space”, *Annals of Mathematics*, **38**(2), (1937), 787-793.
- [32] I. J. Schoenberg, “Metric spaces and positive definite functions”, *Transactions of the American Mathematical Society*, **44**, (1938), 522-536.
- [33] S. Semmes, “Bilipschitz embeddings of metric spaces into Euclidean spaces”, *Publ. Mat.*, **43**, (1999), no. 2, 571-653.
- [34] H. W. Watson, F. Galton, “On the Probability of the Extinction of Families”, *Journal of the Anthropological Institute of Great Britain*, **4**, (1875), 138-144.

JEAN V. BELLISSARD, GEORGIA INSTITUTE OF TECHNOLOGY, SCHOOL OF MATHEMATICS, ATLANTA GA 30332-0160

E-mail address: `jeanbel@math.gatech.edu`

ANTOINE JULIEN, DEPARTMENT OF MATHEMATICAL SCIENCES, NTNU, NO-7491, TRONDHEIM, NORWAY

E-mail address: `antoine.julien@math.ntnu.no`