



HAL
open science

Estimates of Kolmogorov Complexity in Approximating Cantor Sets

C Bonanno, J.-R Chazottes, P. Collet

► **To cite this version:**

C Bonanno, J.-R Chazottes, P. Collet. Estimates of Kolmogorov Complexity in Approximating Cantor Sets. *Nonlinearity*, 2011. hal-02899843

HAL Id: hal-02899843

<https://hal.science/hal-02899843v1>

Submitted on 15 Jul 2020

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

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

Estimates of Kolmogorov Complexity in Approximating Cantor Sets

C. Bonanno

*Dipartimento di Matematica Applicata
Università di Pisa
56127 Pisa, Italy*

J.-R. Chazottes, P. Collet

*Centre de Physique Théorique
Ecole polytechnique, CNRS
91128 Palaiseau, France*

September 8, 2010

Abstract

Our aim is to quantify how complex is a Cantor set as we approximate it better and better. We formalize this by asking what is the shortest program, running on a universal Turing machine, which produces this set at the precision ε in the sense of Hausdorff distance. This is the Kolmogorov complexity of the approximated Cantor set, that we call the “ ε -distortion complexity”. How does this quantity behave as ε tends to 0? And, moreover, how does this behaviour relates to other characteristics of the Cantor set? This is the subject of the present work: we estimate this quantity for several types of Cantor sets on the line generated by iterated function systems (IFS's) and exhibit very different behaviours. For instance, the ε -distortion complexity of a C^k Cantor set is proven to behave like $\varepsilon^{-D/k}$, where D is its box counting dimension.

keywords: iterated function system, random Cantor sets, scaling function, C^k Cantor sets, box counting dimension.

1 Introduction and informal statement of results

In this article, our aim is to quantify the complexity of fractal sets from the point of view of Kolmogorov complexity or Algorithmic Information Content [9, 13, 15]. If we want to draw a fractal set on a computer, we will approximate it by a finite set of points, within a precision ε . We can ask for the shortest program performing this. A natural question is then: how

does the length of such programs behave as ε gets smaller and smaller? More formally, this means that we are interested in the shortest programs, running on a universal Turing machine, which produce this set at the precision ε in the sense of Hausdorff distance. In other words, we consider the Kolmogorov complexity of the approximated set and want to look for its behaviour when ε tends to 0. A basic question is for example whether this complexity depends on the dimension of the set.

Fractal sets appear in many places. Examples include the middle third Cantor set, the Sierpiński triangle, the Koch snowflake; the graphs of Weierstrass functions and Brownian motion; strange attractors of dynamical systems (the Hénon attractor, for instance); Julia and Mandelbrot sets; etc. We refer to, *e.g.*, [17, 3, 2, 11, 18] for different kinds of accounts on fractals. A convenient way to generate many types of fractal sets is to use iterated function systems (IFS for short) [2].

In this article, we consider Cantor sets generated by different kinds of IFS and investigate the behaviour of the Kolmogorov complexity of their approximations in the above sense. For convenience, let us call it the “ ε -distortion complexity”. Our results can be informally summarised as follows.

Informal statement of our results. We first consider IFS’s with polynomial contraction maps and obtain the upper bound $const \times \log(\varepsilon^{-1})$ for the ε -distortion complexity of the generated Cantor set, where the (finite) constant may depend on the polynomials. We can produce “many” polynomial IFS’s with a lower bound of the same order using a probabilistic construction. This shows that our upper bound is the best possible. It turns out that some particular Cantor sets like the usual middle third Cantor set are of much lower complexity.

For analytic IFS’s, we obtain the upper bound $const \times (\log(\varepsilon^{-1}))^2$. We do not have a lower bound in this case (except of course the same lower bound as for IFS’s with polynomial maps). However, as we shall see, the upper bound is more than enough to discriminate clearly between the analytic and the C^k case on the behaviour in ε of the ε -distortion complexity.

Next we consider random central Cantor sets produced by affine IFS’s, for which the contraction rate is chosen at random at each step of the construction. In this case, we get the upper bound $const \times (\log(\varepsilon^{-1}))^2$ and the lower bound $const \times (\log(\varepsilon^{-1}))^{2-\delta}$, for any $\delta > 0$, for almost all such Cantor sets (where the constant in our bound depends on δ and tends to 0 when $\delta \rightarrow 0$).

Finally, we consider C^k IFS’s. Contrarily to the previous cases, the leading, asymptotic behaviour of the ε -distortion complexity depends on the box counting dimension D of the generated Cantor set. Indeed, we obtain the upper bound $const \times \varepsilon^{-\frac{D}{k}-\delta}$, for any $\delta > 0$ (where the constant in our bound depends on δ and blows up when $\delta \rightarrow 0$). We then construct “many” C^k (random) Cantor sets with a lower bound $const \times \varepsilon^{-\frac{D}{k}+\delta}$ (for

any $\delta > 0$), by constructing their scaling function [20]. This shows that our upper bound is the best possible.

Organization of the article. In Section 2 we first define the Kolmogorov complexity of an approximated compact set in the sense of Hausdorff distance. Next, we define the kinds of IFS's we will consider. Section 3 contains our results and Section 4 gathers all the proofs.

We close this introduction by mentioning some works related to ours, though of different nature. One related research stream is about approximation of points in metric spaces. The first ideas on this field date back to the theory of ε -capacity and ε -entropy introduced in [14], and have been developed in [1] in the case of space of functions and approximation of graphs. For a more recent approach using the ideas of Kolmogorov complexity see [12]. The space of all closed subsets of \mathbb{R}^d is a metric space with the Hausdorff distance, hence our results could be interpreted as estimates of the complexity of approximation of Cantor sets in this big metric space. However we are able to distinguish different behaviours of the approximation complexity depending on the nature of the Cantor set, hence giving a deeper insight in the specific problem with respect to the very general approach of approximation in metric spaces.

A second subject is the study of complexity of fractals. Let us first mention that the case of sets reduced to one point on the line was investigated in [8] where in particular the Hausdorff dimension of the set of reals with given asymptotic complexity is computed. For graphs of functions, from the point of view of determining the values of a function at given precision, see [1]. Let us mention that another notion of complexity consists in asking about the smallest execution time of the programs generating a given set with ε -precision, in the sense of Hausdorff distance [21, 7]. Finally, the geometric properties of fractals such as the similarity dimension have been used to characterise the information dimension of points of a fractal [16].

Finally we also mention the book of G. D. Birkhoff [6].

2 Definitions

2.1 Kolmogorov complexity of approximating a set: the ε -distortion complexity

The following definition makes precise what we mean by the Kolmogorov complexity of an “approximated” compact set $\mathcal{C} \subset \mathbb{R}^d$.

Let U denote a Universal Turing machine and consider the set of all binary programs which produce on U a set $w = \{w_1, \dots, w_d\}$ of d finite strings. If we interpret the strings w_i as the co-ordinates of a point $q \in \mathbb{Q}^d$, that is a point with rational co-ordinates, then we can think of the machine

U to produce subsets in \mathbb{R}^d which are finite union of points with rational coordinates. These subsets are the computable “approximations” of a compact set $\mathcal{C} \subset \mathbb{R}^d$ that we consider.

Definition 2.1. *The ε -distortion complexity of a compact set $\mathcal{C} \subset \mathbb{R}^d$ at precision $\varepsilon > 0$, which we call the ε -distortion complexity of \mathcal{C} , is defined by*

$$\Delta(\mathcal{C}, \varepsilon) = \min \{ \ell(\mathbf{P}) : d_{\text{H}}(\mathbf{C}(\mathbf{P}), \mathcal{C}) < \varepsilon \} , \quad (1)$$

where the minimum is taken over all binary programs $\mathbf{P} \in \{0, 1\}^*$ running on a universal Turing machine U , which produce a finite subset $\mathbf{C}(\mathbf{P}) \in \mathbb{R}^d$; $\ell(\mathbf{P})$ is the program length and stop; d_{H} denotes the Hausdorff distance.

Notice that, because of the compactness of \mathcal{C} , we can use a minimum in the above definition, which always leads to a finite number.

We are interested in the behaviour of $\Delta(\mathcal{C}, \varepsilon)$ when ε tends to zero. Note that this is a monotone decreasing function.

For the reader’s convenience, we recall that the Hausdorff distance d_{H} between two closed subsets F_1, F_2 of a metric space with metric d is given by (see, e.g., [18, 2])

$$d_{\text{H}}(F_1, F_2) = \max \left\{ \sup_{x_2 \in F_2} d(x_2, F_1), \sup_{x_1 \in F_1} d(x_1, F_2) \right\}.$$

2.2 Iterated function systems

We now recall the definition of Cantor sets generated by several types of iterated function systems (IFS for short) [2]. For the sake of simplicity, we restrict ourselves to Cantor sets in the unit interval $[0, 1] \subset \mathbb{R}$, although several results can be easily generalised to arbitrary finite dimension.

Let $A = [0, 1]$ and let \mathcal{I} be a finite set of indices with at least two elements. An (hyperbolic) *Iterated Function System* is a collection

$$\{ \phi_i : A \rightarrow A : i \in \mathcal{I} \}$$

of contractions on A with uniform contraction rate (Lipschitz constant) $\rho \in (0, 1)$, and such that $\phi_i(A) \cap \phi_j(A) = \emptyset$ for $i \neq j$. We shall only consider hyperbolic iterated function systems with injective contractions, IHIFS for short. Although some of our results apply in more general cases, this is a class of IFS’s for which the theory of scaling functions is known to hold (see below).

For any infinite word $\omega \in \mathcal{I}^\infty$ and for any $n \in \mathbb{N}$, let $\omega_1^n \in \mathcal{I}^n$ denote the prefix of length n given by the first n symbols of ω , and let

$$\phi_{\omega_1^n} := \phi_{\omega_n} \circ \phi_{\omega_{n-1}} \circ \cdots \circ \phi_{\omega_1}. \quad (2)$$

The map $\pi : \mathcal{I}^\infty \rightarrow A$ defined by $\omega \mapsto \pi(\omega) := \bigcap_{n=0}^\infty \phi_{\omega_1^n}(A)$ is continuous (in product topology) and, since

$$\text{diam} \left(\phi_{\omega_1^n}(A) \right) \leq \rho^n \text{diam} (A),$$

$\pi(\omega)$ is a point in A for all $\omega \in \mathcal{I}^\infty$. The set

$$\mathcal{C} := \pi(\mathcal{I}^\infty) = \bigcup_{\omega \in \mathcal{I}^\infty} \bigcap_{n=0}^\infty \phi_{\omega_1^n}(A)$$

is a Cantor set and satisfies

$$\mathcal{C} = \bigcup_{i \in \mathcal{I}} \phi_i(\mathcal{C}). \quad (3)$$

The classical examples of Cantor sets are the middle $\frac{1}{\rho}$ -th Cantor sets in the unit interval ($\rho = \frac{1}{3}$ gives the usual middle thirds Cantor set). They can be thought of as generated by IHIFS with affine contractions ϕ_0 and ϕ_1 , and contraction rate ρ .

Random affine IHIFS. The additional feature of a random IHIFS is that, at each step of the construction, we consider a random choice for the contraction rate. For the definition we follow [4].

Let us consider a family $(\lambda_k)_{k \in \mathbb{N}}$ of independent identically distributed random variables with values in the interval $(0, 1)$. To each sequence $\lambda = (\lambda_k)_{k \in \mathbb{N}}$ we associate a Cantor set \mathcal{C}_λ in the following way. Let $C_\lambda^0 := [0, 1]$. We define

$$J_1^1(\lambda) := \left[0, \frac{\lambda_1}{2} \right], \quad J_2^1(\lambda) := \left[1 - \frac{\lambda_1}{2}, 1 \right], \quad \text{and} \quad C_\lambda^1 := J_1^1(\lambda) \cup J_2^1(\lambda).$$

In words, C_λ^1 is obtained by removing the central interval of length $(1 - \lambda_1)$ from C_λ^0 . At the $(k + 1)$ -st step, we delete from each interval $J_i^k(\lambda)$, $i = 1, \dots, 2^k$, the central interval of length $(1 - \lambda_{k+1})$, obtaining 2^{k+1} intervals $J_i^{k+1}(\lambda)$, $i = 1, \dots, 2^{k+1}$, such that

$$|J_i^{k+1}(\lambda)| = \frac{1}{2^{k+1}} \prod_{h=1}^{k+1} \lambda_h \quad \forall i = 1, \dots, 2^{k+1}. \quad (4)$$

Then we define

$$C_\lambda^{k+1} := \bigcup_{i=1}^{2^{k+1}} J_i^{k+1}(\lambda).$$

We call *random central Cantor set* the set

$$\mathcal{C}_\lambda := \bigcap_{k=0}^\infty C_\lambda^k.$$

We remark that by construction the boundary points $\partial J_i^k(\lambda)$ of all the intervals $J_i^k(\lambda)$ are contained in \mathcal{C}_λ .

C^k Cantor sets in the line and scaling functions. We now consider Cantor sets with a “differentiable structure”. Their construction is based on the so-called “scaling function” ([20],[19], [5]). For the reader’s convenience and later use, we recall the definition of a differentiable Cantor set in the line and the construction of the scaling function $S_\mathcal{C}$ of a C^k Cantor set \mathcal{C} . In the sequel we fix $\mathcal{I} = \{0, 1\}$ for simplicity, however the argument works for any finite alphabet \mathcal{I} . For a word $\omega_1^n \in \mathcal{I}^n$ we let

$$J_{\omega_1^n} := \phi_{\omega_1^n}([0, 1]) = \phi_{\omega_n} \circ \phi_{\omega_{n-1}} \circ \cdots \circ \phi_{\omega_1}([0, 1]) .$$

Then by definition $\phi_i(J_{\omega_1^n}) = J_{\omega_1^n i}$ for any $i \in \mathcal{I}$, where $\omega_1^n i = \omega_1 \omega_2 \dots \omega_n i$, and it holds

$$J_{\omega_1^n} \subset J_{\omega_2^n} \subset \cdots \subset J_{\omega_{n-1}^n} \subset J_{\omega_n} . \quad (5)$$

Moreover we assume that the continuous map $\pi : \mathcal{I}^\infty \rightarrow [0, 1]$ is order-preserving, and let $\mathcal{C} := \pi(\mathcal{I}^\infty)$.

Let now $\sigma : \mathcal{I}^\infty \rightarrow \mathcal{I}^\infty$ denote the standard shift-map and $\{\sigma_i^{-1}\}_{i \in \mathcal{I}}$ be the set of right-inverses of σ , that is $\sigma_i^{-1}(\omega) = i\omega$. Then the map π induces a shift-map $f : \mathcal{C} \rightarrow \mathcal{C}$ with right-inverses $\{f_i^{-1}\}_{i \in \mathcal{I}}$. The Cantor set \mathcal{C} is said to be of class C^k if each of the right-inverses f_i^{-1} has a C^k extension to $A = [0, 1]$ which is a contraction.

The scaling function instead describes the contraction rates in the inclusions (5). For a word $\omega_1^n \in \mathcal{I}^n$ we define $\tilde{S}_\mathcal{C}(\omega_1^n) \in (0, 1)^2$. The two components of $\tilde{S}_\mathcal{C}(\omega_1^n)$ are the rates of contractions

$$(\tilde{S}_\mathcal{C}(\omega_1^n))_i = \frac{|J_{i\omega_1^n}|}{|J_{\omega_1^n}|} \quad i = 0, 1$$

where $|J|$ denotes the length of the interval J . The length of the gap between the two intervals $J_{0\omega_1^n}$ and $J_{1\omega_1^n}$ in $J_{\omega_1^n}$ can be reconstructed from these data. The scaling function is defined to be the function

$$S_\mathcal{C} : \mathcal{I}^\infty \rightarrow (0, 1)^2$$

given by

$$S_\mathcal{C}(\omega) := \lim_{n \rightarrow \infty} \tilde{S}_\mathcal{C}(\omega_1^n) .$$

We refer to [20] for the proof of the existence of this limit.

By definition one has

$$|J_{\omega_1^n}| = \prod_{j=1}^{n-1} \left(\tilde{S}_\mathcal{C}(\omega_{j+1}^n) \right)_{\omega_j} .$$

By using the scaling function we can introduce a distance $d_S(\omega, \tilde{\omega})$ on \mathcal{I}^∞ in the following way. For two sequences $\omega, \tilde{\omega} \in \mathcal{I}^\infty$, let $\omega \cap \tilde{\omega}$ denote their longest common prefix, and let $|\omega \cap \tilde{\omega}|$ denote its length. Then we let

$$d_S(\omega, \tilde{\omega}) := \sup_{\alpha \in \mathcal{I}^\infty} \prod_{j=1}^{n=|\omega \cap \tilde{\omega}|} (S_{\mathcal{C}}(\omega_{j+1}^n \alpha))_{\omega_j} . \quad (6)$$

Then there exists a constant $H > 0$ such that for any $\omega, \tilde{\omega} \in \mathcal{I}^\infty$ it holds

$$\frac{1}{H} \leq \frac{|J_{\omega \cap \tilde{\omega}}|}{d_S(\omega, \tilde{\omega})} \leq H .$$

Relations between the properties of the scaling function of a C^k central Cantor set and the differentiability of the IHIFS generating this Cantor set have been studied in [19] in the case $k > 1$. In particular, Main Theorem [19, pag. 406] introduces the following characterisation of a C^k Cantor set in the line. Let $A(\omega_1^n)$ denote the set of the four boundary points of the intervals $(J_{i\omega_1^n})_{i \in \mathcal{I}}$. A scaling function $S_{\mathcal{C}}$ generates a C^k Cantor set \mathcal{C} if and only if for any $n \in \mathbb{N}$ there are diffeomorphisms from $A(\omega_1^n)$ into $A(\tilde{\omega}_1^n)$, for any $\omega_1^n \neq \tilde{\omega}_1^n \in \mathcal{I}^n$, with derivatives bounded by a constant $C(\omega_1^n, \tilde{\omega}_1^n)$ which satisfies

$$C(\omega_1^n, \tilde{\omega}_1^n) = C d_S(\omega_1^n \alpha, \tilde{\omega}_1^n \alpha)^{k-1} \quad (7)$$

where C does not depend on n and from the definition of d_S the right hand side is independent on $\alpha \in \mathcal{I}^\infty$.

3 Results

The next four theorems give the behaviour, as $\varepsilon \rightarrow 0$, of the ε -distortion complexity (1), for different types of IHIFS, namely, polynomial IHIFS, real analytic IHIFS, random affine IHIFS and C^k IHIFS. We will use the following notations.

Notations. *In the sequel we write $f \asymp g$ if there are two positive constants C_1 and C_2 such that for any $\varepsilon > 0$ small enough*

$$C_1 f(\varepsilon) \leq g(\varepsilon) \leq C_2 f(\varepsilon).$$

We write $f \preccurlyeq g$ if there is a positive constant C such that for any $\varepsilon > 0$ small enough

$$f(\varepsilon) \leq C g(\varepsilon).$$

Our first result deals with polynomial IHIFS's.

Theorem 3.1. *Let \mathcal{C} be a Cantor set generated by an IHIFS with polynomial functions. Then*

$$\Delta(\mathcal{C}, \varepsilon) \asymp \log(\varepsilon^{-1}). \quad (8)$$

Moreover, for any $\delta > 0$, there exist (many) polynomial IHIFS's such that the generated Cantor set satisfies

$$(1 - \delta) \log(\varepsilon^{-1}) \leq \Delta(\mathcal{C}, \varepsilon). \quad (9)$$

Remark 3.1. *A more precise upper bound follows easily from the proof, namely*

$$\Delta(\mathcal{C}, \varepsilon) \leq \left(\sum_{i \in \mathcal{I}} (1 + \deg \phi_i) \right) \log(\varepsilon^{-1}) + o(\log(\varepsilon^{-1})).$$

A more precise lower bound of the same kind can also be obtained for a large class of Cantor sets which are generated by a set of full measure of some random polynomial IHIFS's.

Remark 3.2. *Obviously, the classical middle third Cantor set (or more generally any middle $\frac{1}{\rho}$ -th set with ρ a rational number) has a much smaller complexity.*

Our next result is about real analytic IHIFS's.

Theorem 3.2. *Let \mathcal{C} be a Cantor set generated by an IHIFS with real analytic functions. Then*

$$\Delta(\mathcal{C}, \varepsilon) \asymp (\log(\varepsilon^{-1}))^2.$$

We do not have a lower bound in this case (except of course the polynomial lower bound). However, as we shall see, the upper bound is more than enough to discriminate clearly between the analytic and the C^k case on the behaviour in ε of the ε -distortion complexity.

The next theorem states that random central Cantor sets need more information than those generated by polynomial IHIFS's.

Theorem 3.3. *Let \mathcal{C}_λ be a random central Cantor set as described above. Then, for any $\lambda \in (0, 1)^\mathbb{N}$,*

$$\Delta(\mathcal{C}_\lambda, \varepsilon) \asymp (\log(\varepsilon^{-1}))^2. \quad (10)$$

Moreover, let us assume that the common distribution of the i.i.d. random variables (λ_k) is absolutely continuous, with a density $f(x)$ bounded above and below away from zero. Then, for any $\delta > 0$, we have

$$(\log(\varepsilon^{-1}))^{2-\delta} \asymp \Delta(\mathcal{C}_\lambda, \varepsilon) \quad (11)$$

for almost every $\lambda \in (0, 1)^\mathbb{N}$.

We now consider C^k Cantor sets with a differentiable structure.

Theorem 3.4. *Let $k > 1$. For any $\delta > 0$, for any C^k Cantor set \mathcal{C} with box counting dimension D , we have*

$$\Delta(\mathcal{C}, \varepsilon) \asymp \varepsilon^{-\frac{D}{k}-\delta}. \quad (12)$$

Moreover, for any $\delta > 0$, there exist (many) C^k central Cantor sets \mathcal{C} with box counting dimension, at most $D + \delta$, such that

$$\varepsilon^{-\frac{D}{k}+\delta} \asymp \Delta(\mathcal{C}, \varepsilon). \quad (13)$$

We emphasise that in this case the asymptotic behaviour of the ε -distortion complexity $\Delta(\mathcal{C}, \varepsilon)$, when ε tends to zero, depends in general on the regularity k of the set and, contrarily to the previous cases, it also depends on its box counting dimension D .

4 Proofs

The following two simple lemmas will be used repeatedly in the proofs hereafter. We leave their elementary proof to the reader.

Lemma 4.1. *Let F and F' be closed subsets of A . Let $I = [a, b]$ and $I' = [a', b']$ be closed sub-intervals of A . Let $H = [c, d]$ and $H' = [c', d']$ be closed subsets of $\overset{\circ}{I}$ and $\overset{\circ}{I}'$ respectively. Assume that $\partial H \subset F$, $\partial H' \subset F'$, $F \cap \overset{\circ}{H} = \emptyset$ and $F' \cap \overset{\circ}{H}' = \emptyset$. Moreover assume that there exists $\varepsilon > 0$ such that $|a - a'| \leq \varepsilon$, $|b - b'| \leq \varepsilon$, $|c - d| > 2\varepsilon$, $|c' - d'| > 2\varepsilon$ and $\max\{|c - c'|, |d - d'|\} > \varepsilon$, then $d_H(F, F') > \varepsilon$.*

In the sequel, this lemma will be used to show that two Cantor sets (F and F') are at Hausdorff distance larger than ε , H and H' playing the role of holes in the Cantor sets.

Lemma 4.2. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let \mathcal{C} be a measurable map from (Ω, \mathcal{A}) to the set of closed subsets of A equipped with the Borel σ -algebra induced by the Hausdorff metric. Let $(a_k)_k$ be a positive, increasing, diverging sequence. Assume that for any integer k there exists a sequence $(V_{k,j})_{1 \leq j \leq 2^{a_k}}$ of measurable subsets of Ω such that*

$$\left\{ \omega : \Delta(\mathcal{C}(\omega), 2^{-k}) < a_k \right\} \subset \bigcup_{j=1}^{2^{a_k}} V_{k,j}$$

and

$$\sum_k \sum_{j=1}^{2^{a_k}} \mathbb{P}(V_{k,j}) < \infty.$$

Then, for \mathbb{P} -almost every ω , $\Delta(\mathcal{C}(\omega), 2^{-k}) \geq a_k$ for any k large enough (depending on ω).

All proofs have the same structure: first an upper bound is obtained by explicitly constructing an approximation and estimating the complexity; second a lower bound is constructed using the two previous lemmas.

4.1 Proof of Theorem 3.1

Let N be the largest degree of the polynomial functions $\{\phi_i\}$, then we can write

$$\phi_i(x) = \sum_{0 \leq \alpha \leq N} c_{i,\alpha} x^\alpha \quad \forall i \in \mathcal{I}$$

with coefficients $c_{i,\alpha} \in \mathbb{R}$. We now show how to construct a program P approximating \mathcal{C} within Hausdorff distance ε .

Let ε be fixed and K a constant to be specified later on. Let us define $\varepsilon' = \frac{\varepsilon}{K}$. We construct polynomials

$$\tilde{\phi}_i(x) := \sum_{0 \leq \alpha \leq N} \tilde{c}_{i,\alpha} x^\alpha \quad \forall i \in \mathcal{I}$$

with coefficients satisfying

$$|c_{i,\alpha} - \tilde{c}_{i,\alpha}| < \varepsilon' \quad \forall i \in \mathcal{I} \quad \forall 0 \leq \alpha \leq N \quad (14)$$

such that the $\tilde{\phi}_i$'s are injective contractions on $A(= [0, 1])$ with uniform contraction rate $\tilde{\rho} \in (0, 1)$. For any $\omega_1^n \in \mathcal{I}^n$ we construct the composition $\tilde{\phi}_{\omega_1^n}$ as in (2).

We first show that for any bounded set $B \subset \mathbb{R}$ such that $\phi_i(B) \subset B$ for all $i \in \mathcal{I}$ with the same contraction rate ρ , and $\tilde{\phi}_i(B) \subset B$ for all $i \in \mathcal{I}$, we have for all $n \in \mathbb{N}$

$$d_H(\phi_{\omega_1^n}(B), \tilde{\phi}_{\omega_1^n}(B)) < \varepsilon' \frac{1 - \rho^n}{1 - \rho} \sup_{x \in B} \left| \sum_{0 \leq \alpha \leq N} x^\alpha \right| \quad \forall \omega_1^n \in \mathcal{I}^n. \quad (15)$$

The proof is by induction. By (14) and definition of d_H , one immediately gets

$$d_H(\phi_i(B), \tilde{\phi}_i(B)) < \varepsilon' \sup_{x \in B} \left| \sum_{0 \leq \alpha \leq N} x^\alpha \right| \quad \forall i \in \mathcal{I}.$$

The inductive step is obtained by using the triangle inequality for d_H . By the first step we have

$$d_H\left(\phi_{\omega_n}(\tilde{\phi}_{\omega_1^{n-1}}(B)), \tilde{\phi}_{\omega_n}(\tilde{\phi}_{\omega_1^{n-1}}(B))\right) < \varepsilon' \sup_{x \in B} \left| \sum_{0 \leq \alpha \leq N} x^\alpha \right|$$

where we have used $\tilde{\phi}_{\omega_1^{n-1}}(B) \subset B$. Moreover, by using the contraction rate ρ , we get

$$\begin{aligned} d_{\mathbb{H}}\left(\phi_{\omega_n}(\phi_{\omega_1^{n-1}}(B)), \phi_{\omega_n}(\tilde{\phi}_{\omega_1^{n-1}}(B))\right) &< \rho d_{\mathbb{H}}(\phi_{\omega_1^{n-1}}(B), \tilde{\phi}_{\omega_1^{n-1}}(B)) \\ &< \varepsilon' \rho \frac{1 - \rho^{n-1}}{1 - \rho} \sup_{x \in B} \left| \sum_{0 \leq \alpha \leq N} x^\alpha \right| \end{aligned}$$

where the last inequality is the $(n - 1)$ -th step of the induction. Hence the triangle inequality implies that

$$d_{\mathbb{H}}(\phi_{\omega_1^n}(B), \tilde{\phi}_{\omega_1^n}(B)) < \varepsilon' \left(1 + \rho \frac{1 - \rho^{n-1}}{1 - \rho}\right) \sup_{x \in B} \left| \sum_{0 \leq \alpha \leq N} x^\alpha \right|.$$

This finishes the proof of (15).

Let us choose $\bar{n} \in \mathbb{N}$ such that $\rho^{\bar{n}} < \varepsilon'$ and $\tilde{\rho}^{\bar{n}} < \varepsilon'$. For this fixed \bar{n} , let $V := \{0, 1\} = \partial[0, 1]$ and define

$$\mathbf{c} := \bigcup_{\omega_1^{\bar{n}} \in \mathcal{I}^{\bar{n}}} \tilde{\phi}_{\omega_1^{\bar{n}}}(V).$$

We now prove that $d_{\mathbb{H}}(\mathcal{C}, \mathbf{c}) < \varepsilon$. Let us consider $x \in \mathcal{C}$ and $y \in \mathbf{c}$. By (3) there exists $z \in \mathcal{C}$ such that $x = \phi_{\omega_1^{\bar{n}}(x)}(z)$ for a given sequence $\omega_1^{\bar{n}}(x) \in \mathcal{I}^{\bar{n}}$. Hence

$$\begin{aligned} d(x, y) &= d(\phi_{\omega_1^{\bar{n}}(x)}(z), y) \leq \\ &d_{\mathbb{H}}(\phi_{\omega_1^{\bar{n}}(x)}(z), \phi_{\omega_1^{\bar{n}}(x)}(V)) + d_{\mathbb{H}}(\phi_{\omega_1^{\bar{n}}(x)}(V), \tilde{\phi}_{\omega_1^{\bar{n}}(x)}(V)) + d_{\mathbb{H}}(\tilde{\phi}_{\omega_1^{\bar{n}}(x)}(V), y). \end{aligned} \tag{16}$$

For the first term we use the contraction properties to get

$$d_{\mathbb{H}}(\phi_{\omega_1^{\bar{n}}(x)}(z), \phi_{\omega_1^{\bar{n}}(x)}(V)) < \rho^{\bar{n}} \text{diam}(A) < \varepsilon'.$$

By (15), for the second term we have

$$d_{\mathbb{H}}(\phi_{\omega_1^{\bar{n}}(x)}(V), \tilde{\phi}_{\omega_1^{\bar{n}}(x)}(V)) < \frac{\varepsilon'}{1 - \rho} (N + 1).$$

If we take

$$K = 1 + \frac{N + 1}{1 - \rho}$$

then

$$d(x, y) \leq \varepsilon' K + d_{\mathbb{H}}(\tilde{\phi}_{\omega_1^{\bar{n}}(x)}(V), y) = \varepsilon + d_{\mathbb{H}}(\tilde{\phi}_{\omega_1^{\bar{n}}(x)}(V), y).$$

Choosing $y \in \tilde{\phi}_{\omega_1^{\bar{n}}(x)}(V) \subset \mathbf{C}$, we have

$$d_{\mathbf{H}}(\tilde{\phi}_{\omega_1^{\bar{n}}(x)}(V), y) = 0,$$

hence $d(x, y) < \varepsilon$. Therefore

$$\sup_{x \in \mathcal{C}} d(x, \mathbf{C}) < \varepsilon.$$

On the other hand for a given $\omega_1^{\bar{n}} \in \mathcal{I}^{\bar{n}}$ and $y \in \tilde{\phi}_{\omega_1^{\bar{n}}}(V)$, take $x \in \phi_{\omega_1^{\bar{n}}}(A) \cap \mathcal{C}$, noticing that this set is not empty. Then we deduce that

$$\sup_{y \in \mathbf{C}} d(y, \mathcal{C}) < \varepsilon.$$

Hence $d_{\mathbf{H}}(\mathcal{C}, \mathbf{C}) < \varepsilon$.

Let us define the program \mathbf{P} that contains the numbers ε , ρ and K , and such that it specifies the coefficients $\{\tilde{c}_{i,\alpha}\}$, computes \bar{n} and makes the computation of the $\tilde{\phi}_i(V)$'s. The binary length $\ell(\mathbf{P})$ satisfies

$$\ell(\mathbf{P}) \preccurlyeq \log((\varepsilon')^{-1}) = \log K + \log(\varepsilon^{-1}).$$

Indeed, ε is specified with $O(\log \varepsilon^{-1})$ bits, and ρ and K do not depend on ε and can be approximated by rational numbers. The coefficients $\{\tilde{c}_{i,\alpha}\}$ are approximations of the $\{c_{i,\alpha}\}$ with precision ε' , hence we can choose them as rational numbers requiring only $O(\log(\varepsilon')^{-1})$ bits of information. Finally the information for the computation of \bar{n} and \mathbf{C} needs only $O(1)$ bits of information. Hence this proves (8).

We now prove (9). Let $\mathcal{I} = \{0, 1\}$ and define $\phi_1(x) = bx$, $\phi_2(x) = 1 - bx$ for $b \in (0, 1/2)$. (Note that these are injective functions.) We denote by \mathcal{C}_b the Cantor set generated by the IHIFS $\{\phi_1, \phi_2\}$. To be in the context of Lemma 4.2, we take b at random according to the uniform distribution on the interval $(1/4, 1/3)$. We restrict the possible values of b to ensure that the middle hole is large enough so that an obviously simplified version of Lemma 4.1 applies.

For a fixed $\delta \in (0, 1)$, define $a_k := (1 - \delta)k$. For any k , there are at most 2^{a_k} different binary programs $(\mathbf{P}_j)_{1 \leq j \leq 2^{a_k}}$ of length $a_k - 1$, which generate at most 2^{a_k} different sets $\mathbf{C}_j := \mathbf{C}(\mathbf{P}_j)$. We define

$$V_{k,j} := \left\{ b : d_{\mathbf{H}}(\mathcal{C}_b, \mathbf{C}_j) < 2^{-k} \right\}.$$

Then

$$\left\{ b : \Delta(\mathcal{C}_b, 2^{-k}) < a_k \right\} \subset \bigcup_{j=1}^{2^{a_k}} V_{k,j}.$$

We now estimate $\mathbb{P}(V_{k,j})$. We denote by $\partial^+ \mathbb{C}_j$ the rightmost point of $\mathbb{C}_j \cap (0, 1/2)$. For k large enough ($k \geq 3$) and for a given j , if $b \in V_{k,j}$ then $d(b, \partial^+ \mathbb{C}_j) < 2^{-k}$ and therefore $\mathbb{P}(V_{k,j}) < 2^{1-k}$. This implies that

$$\sum_k \sum_{j=1}^{2^{a_k}} \mathbb{P}(V_{k,j}) < \sum_k 2^{a_k} 2^{1-k} = \sum_k 2^{1-\delta k} < \infty.$$

The result follows from Lemma 4.2.

Remark 4.1. *Notice that this proof works also in arbitrary finite dimension.*

4.2 Proof of Theorem 3.2

We give the proof in the case that the $\{\phi_i\}$ are analytic functions on an open ball $B(0, R) \subset \mathbb{C}$ of radius $R > 1$. The general case follows by applying the same argument to piecewise polynomial approximations of the $\{\phi_i\}$.

By hypothesis we can write for $z \in A$

$$\phi_i(z) = \sum_{h=0}^{\infty} c_{i,h} z^h \quad \forall i \in \mathcal{I}$$

with coefficients $c_{i,h} \in \mathbb{C}$. By the analyticity of the functions $\{\phi_i\}$ in $B(0, R)$ it follows that

$$\limsup_{h \rightarrow \infty} |c_{i,h}| R^h \leq 1.$$

Hence there exists a real constant r such that

$$|c_{i,h}| R^h \leq r \quad \forall h \geq 0.$$

As above let us denote by $V = [0, 1] = \partial[0, 1]$. We now construct approximations of the analytic functions $\{\phi_i\}$.

Let $\delta \in (0, 1)$ be such that $\delta R > 1$, and, for ε fixed, define

$$\varepsilon' := \frac{r(\delta R - 1)}{R(1 - \delta)} ((1 - \delta)\varepsilon)^{\left(\frac{\log R}{\log \frac{1}{\delta}}\right)}. \quad (17)$$

Let N be the smallest integer satisfying

$$\sum_{h=N}^{\infty} \delta^h = \frac{\delta^N}{1 - \delta} < (1 - \rho) \frac{\varepsilon}{4r}. \quad (18)$$

Hence we construct the polynomials

$$\tilde{\phi}_i(z) := \sum_{h=0}^{N-1} \tilde{c}_{i,h} z^h \quad \forall i \in \mathcal{I}$$

with coefficients $\tilde{c}_{i,h} \in \mathbb{C}$ such that

$$|c_{i,h} - \tilde{c}_{i,h}| < \varepsilon' \quad \forall i \in \mathcal{I} \quad \forall h = 0, \dots, N-1 \quad (19)$$

and the $\tilde{\phi}_i$'s are contractions functions on A with $\tilde{\phi}_i(A) \subset A$ for all $i \in \mathcal{I}$, and uniform contraction rate $\tilde{\rho} \in (0, 1)$.

Let us choose $\bar{n} \in \mathbb{N}$ such that $\rho^{\bar{n}} < \frac{\varepsilon}{4R}$ and $\tilde{\rho}^{\bar{n}} < \frac{\varepsilon}{4R}$. For this fixed \bar{n} , we define

$$\mathbf{c} := \bigcup_{\omega_1^{\bar{n}} \in \mathcal{I}^{\bar{n}}} \tilde{\phi}_{\omega_1^{\bar{n}}}(V).$$

We now prove that $d_{\mathbb{H}}(\mathcal{C}, \mathbf{c}) < \varepsilon$. The proof will follow again by using (16) and the analogue of (15).

For any bounded set $B \subset B(0, \delta R)$ such that $\phi_i(B) \subset B$ and $\tilde{\phi}_i(B) \subset B$, we have for all $n \in \mathbb{N}$

$$d_{\mathbb{H}}(\phi_{\omega_1^n}(B), \tilde{\phi}_{\omega_1^n}(B)) < \varepsilon \frac{1 - \rho^n}{2} \quad \forall \omega_1^n \in \mathcal{I}^n. \quad (20)$$

The proof of (20) is by induction as the proof of (15). We only show the first step. By using (19) and (18), we obtain for all $z \in B(0, \delta R)$

$$\begin{aligned} |\phi_i(z) - \tilde{\phi}_i(z)| &= \left| \sum_{h=0}^{N-1} (c_{i,h} - \tilde{c}_{i,h}) z^h + \sum_{h=N}^{\infty} c_{i,h} z^h \right| \\ &\leq \varepsilon' \sum_{h=0}^{N-1} \delta^h R^h + r \sum_{h=N}^{\infty} \delta^h < \varepsilon' \frac{\delta^N R^N}{R\delta - 1} + (1 - \rho) \frac{\varepsilon}{4r} r \\ &< (1 - \rho) \varepsilon \left(\varepsilon' \frac{R(1 - \delta)((1 - \delta)\varepsilon)^{\left(-\frac{\log R}{\log \delta - 1}\right)}}{4r(\delta R - 1)} + \frac{r}{4r} \right) = (1 - \rho) \frac{\varepsilon}{2}, \end{aligned}$$

where in the last equality we have used (17). Hence

$$d_{\mathbb{H}}(\phi_i(B), \tilde{\phi}_i(B)) < \varepsilon \frac{(1 - \rho)}{2} \quad \forall i \in \mathcal{I}.$$

The inductive step is obtained as in the proof of (15) by the triangle inequality for $d_{\mathbb{H}}$.

We now write (16) and, by repeating the argument of the proof of Theorem 3.1 and by (20), we obtain $d_{\mathbb{H}}(\mathcal{C}, \mathbf{c}) < \varepsilon$.

Let us define the program \mathbf{P} which contains the numbers ε , R , r , δ , N and \bar{n} , and such that it specifies the coefficients $\{\tilde{c}_{i,h}\}$ for $h = 0, \dots, N-1$, and makes the computation of the $\tilde{\phi}_i(V)$'s. The binary length $\ell(\mathbf{P})$ satisfies

$$\ell(\mathbf{P}) \preceq N \log((\varepsilon')^{-1}) \preceq (\log(\varepsilon^{-1}))^2.$$

Indeed, ε is specified with $O(\log \varepsilon^{-1})$ bits, and $\bar{n} \asymp \log \varepsilon^{-1}$. Hence it is specified by $o(\log \varepsilon^{-1})$ bits and $N \asymp \log \varepsilon^{-1}$, hence it is specified by $o(\log \varepsilon^{-1})$ bits. The coefficients $\{\tilde{c}_{i,h}\}$ are approximations of the $\{c_{i,h}\}$ with precision ε' , hence we can choose them as rational numbers requiring only $O(\log(\varepsilon')^{-1})$ bits of information, but there are N of them for each function ϕ_i (see (17) and (18)), hence we need $O(N \log(\varepsilon')^{-1})$ bits of information, that is $O((\log(\varepsilon^{-1}))^2)$. Finally R , r and δ do not depend on ε , hence the information for them and for the computation of \mathbf{C} need only $O(1)$ bits of information. Hence

$$\Delta(\mathcal{C}, \varepsilon) \preceq (\log(\varepsilon^{-1}))^2$$

and the theorem is proved.

4.3 Proof of Theorem 3.3

We first prove (10) by constructing an approximation \mathbf{C} of the set \mathcal{C} .

Let us consider a fixed sequence $\lambda \in (0, 1)^{\mathbb{N}}$ and the Cantor set \mathcal{C}_λ . For a fixed ε , let \bar{n} be given by

$$\bar{n} := \min \left\{ n \in \mathbb{N} : \frac{1}{2^n} < \frac{\varepsilon}{2} \right\}. \quad (21)$$

Next, let us consider a sequence $\tilde{\lambda} = (\tilde{\lambda}_k)_{k \in \mathbb{N}} \in (0, 1)^{\mathbb{N}}$ such that $|\lambda_1 - \tilde{\lambda}_1| < \varepsilon$ and

$$|\lambda_k - \tilde{\lambda}_k| < 2^{k-2} \varepsilon \quad \forall k = 2, \dots, \bar{n}. \quad (22)$$

Then we define the approximation \mathbf{C} of \mathcal{C}_λ to be the finite set

$$\mathbf{C} := \partial C_{\tilde{\lambda}}^{\bar{n}} = \bigcup_{i=1}^{2^{\bar{n}}} \partial J_i^{\bar{n}}(\tilde{\lambda})$$

where the sets $C_{\tilde{\lambda}}^{\bar{n}}$ and $J_i^{\bar{n}}(\tilde{\lambda})$ are constructed as specified in Section 2. We now prove that $d_{\text{H}}^{\tilde{\lambda}}(\mathcal{C}_\lambda, \mathbf{C}) < \varepsilon$. To this aim we show that

$$d_{\text{H}}(\partial J_1^k(\lambda), \partial J_1^k(\tilde{\lambda})) < \frac{\varepsilon}{2} \quad (23)$$

for all $k = 1, \dots, \bar{n}$. The same argument applies to all other sets J_i^k . This is enough since it implies that, for any two points $x \in \mathcal{C}_\lambda$ and $y \in \mathbf{C}$ in the analogous intervals (*i.e.*, $x \in J_i^{\bar{n}}(\lambda)$ and $y \in \partial J_i^{\bar{n}}(\tilde{\lambda})$ with the same index $i = 1, \dots, 2^{\bar{n}}$),

$$\begin{aligned} d(x, y) &\leq d(x, \partial J_i^{\bar{n}}(\lambda)) + d_{\text{H}}(\partial J_i^{\bar{n}}(\lambda), \partial J_i^{\bar{n}}(\tilde{\lambda})) + d(\partial J_i^{\bar{n}}(\tilde{\lambda}), y) \\ &< \frac{1}{2^{\bar{n}}} + \frac{\varepsilon}{2} + 0 < \varepsilon, \end{aligned}$$

where for the first term we have used (4) and (21), for the second term we have used (23), and for the third term we have used the definition of \mathbf{C} . Hence $d_{\text{H}}(\mathcal{C}_\lambda, \mathbf{C}) < \varepsilon$. It remains to prove (23). By definition

$$\partial J_1^k(\lambda) = \left\{ 0, \frac{1}{2^k} \prod_{h=1}^k \lambda_h \right\}, \partial J_1^k(\tilde{\lambda}) = \left\{ 0, \frac{1}{2^k} \prod_{h=1}^k \tilde{\lambda}_h \right\}.$$

Hence it is enough to show that

$$d\left(\frac{1}{2^k} \prod_{h=1}^k \lambda_h, \frac{1}{2^k} \prod_{h=1}^k \tilde{\lambda}_h\right) < \frac{\varepsilon}{2} \quad (24)$$

for all $k = 1, \dots, \bar{n}$. By definition of $\tilde{\lambda}$, it holds

$$d\left(\frac{1}{2} \lambda_1, \frac{1}{2} \tilde{\lambda}_1\right) < \frac{\varepsilon}{2}.$$

Assuming that (24) holds for $k - 1$, one gets

$$\begin{aligned} & d\left(\frac{1}{2^k} \prod_{h=1}^k \lambda_h, \frac{1}{2^k} \prod_{h=1}^k \tilde{\lambda}_h\right) \\ & \leq \left| \frac{1}{2^{k-1}} \prod_{h=1}^{k-1} \lambda_h \right| \left| \frac{\lambda_k - \tilde{\lambda}_k}{2} \right| + d\left(\frac{1}{2^{k-1}} \prod_{h=1}^{k-1} \lambda_h, \frac{1}{2^{k-1}} \prod_{h=1}^{k-1} \tilde{\lambda}_h\right) \frac{\tilde{\lambda}_k}{2} \\ & < \frac{1}{2^{k-1}} \frac{2^{k-2} \varepsilon}{2} + \frac{\varepsilon}{4} < \frac{\varepsilon}{2}, \end{aligned}$$

where we have used (22) for the first term, the inductive hypothesis for the second term, and the fact that $\lambda, \tilde{\lambda} \in (0, 1)^{\mathbb{N}}$.

To finish the proof of (10) we have to estimate the length of a binary program \mathbf{P} producing the finite set \mathbf{C} . The program \mathbf{P} must contain the number ε , the information to compute \bar{n} , the contraction factors $(\tilde{\lambda}_k)$ for $k = 1, \dots, \bar{n}$ and the instruction to compute \mathbf{C} . The number of instructions for all the computations are $O(1)$ with respect to ε . The number ε is specified by $O(\log \varepsilon^{-1})$ bits of information and $\bar{n} \approx \log \varepsilon^{-1}$ and each coefficient $\tilde{\lambda}_k$ needs $O(\log \varepsilon^{-1})$ bits of information. Since there are \bar{n} coefficients to be specified, we find

$$\ell(\mathbf{P}) \preceq (\log(\varepsilon^{-1}))^2$$

hence (10) follows.

We now prove (11). First of all we identify a full measure set of “good” λ . By the hypothesis on the density $f(x)$ of the common distribution of the random variables (λ_k) , the following quantity is finite

$$\gamma := \int_0^1 \log(x) f(x) dx < 0.$$

Notice that $\frac{e^\gamma}{2}$ can be interpreted as the typical contraction rate, since products of many i.i.d. random variables λ_k will be involved.

Given any $\eta > 0$ we define, for all $n \in \mathbb{N}$,

$$\Lambda_n := \left\{ \lambda \in (0, 1)^{\mathbb{N}} : \prod_{h=1}^k \lambda_h > e^{k(\gamma-\eta)}, \forall k = \lfloor \sqrt{n} \rfloor, \dots, n \right\}. \quad (25)$$

We remark that, since $\lambda_k \in (0, 1)$ for all $k \geq 1$, if $\lambda \in \Lambda_n$ then

$$\prod_{h=1}^k \lambda_h > e^{\sqrt{n}(\gamma-\eta)} \quad \forall k = 1, \dots, \lfloor \sqrt{n} \rfloor - 1. \quad (26)$$

Lemma 4.3. *Let us denote by Λ_n^c the complement of Λ_n in $(0, 1)^{\mathbb{N}}$, then*

$$\sum_n \mathbb{P}(\Lambda_n^c) < \infty,$$

i.e., almost every $\lambda \in (0, 1)^{\mathbb{N}}$ belongs to Λ_n^c only for finitely many $n \in \mathbb{N}$.

Proof. We use the large deviation principle for independent and identically distributed random variables (see, e.g., [10]). It implies that for any fixed $\eta > 0$ there exists a positive constant C such that

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \mathbb{P} \left\{ \frac{\log(\lambda_1 \times \dots \times \lambda_k)}{k} - \gamma < -\eta \right\} = -C.$$

Hence, for n large enough and for all $0 < C' < C$, we have the estimate

$$\mathbb{P}(\Lambda_n^c) \leq \sum_{k=\lfloor \sqrt{n} \rfloor}^n \mathbb{P} \left\{ \prod_{h=1}^k \lambda_h < e^{k(\gamma-\eta)} \right\} \leq e^{-C' \sqrt{n}} \sum_{k=\lfloor \sqrt{n} \rfloor}^n e^{-C'(k-\sqrt{n})}.$$

Therefore

$$\mathbb{P}(\Lambda_n^c) \leq e^{-C' \sqrt{n}}$$

and the lemma follows by the Borel-Cantelli Lemma. \square

For any ε we define

$$N(\varepsilon) := \min \left\{ n \in \mathbb{N} : (2e^{\eta-\gamma})^n \varepsilon > (\log(\varepsilon^{-1}))^{-2} \right\}. \quad (27)$$

We now consider a subset of $\Lambda_{N(\varepsilon)}$. Let $q \in \mathbb{N}$ and define

$$\Psi_q := \left\{ \lambda \in \Lambda_{N(2^{-q})} : (1 - \lambda_{k+1}) \frac{\prod_{h=1}^k \lambda_h}{2^k} > 2^{2-q}, \forall k = 0, \dots, N(2^{-q}) \right\}.$$

Lemma 4.4. *We have*

$$\sum_q \mathbb{P}(\Psi_q^c) < \infty$$

i.e., almost every $\lambda \in (0, 1)^{\mathbb{N}}$ lies in Ψ_q^c only for finitely many $q \in \mathbb{N}$.

Proof. By Lemma 4.3 and the Borel-Cantelli Lemma, it is enough to prove that

$$\sum_q \mathbb{P}(\Psi_q^c \cap \Lambda_{N(2^{-q})}) < \infty.$$

First observe that if $0 \leq k \leq \sqrt{N(2^{-q})}$ and

$$(1 - \lambda_{k+1}) > 2^{2-q} 2^k (e^{\eta-\gamma})^{\sqrt{N(2^{-q})}}$$

then λ satisfies

$$(1 - \lambda_{k+1}) \frac{\prod_{h=1}^k \lambda_h}{2^k} > 2^{2-q}. \quad (28)$$

Similarly, if $\sqrt{N(2^{-q})} \leq k \leq N(2^{-q})$ and

$$(1 - \lambda_{k+1}) > 2^{2-q} 2^k (e^{\eta-\gamma})^k,$$

then (28) holds. Therefore

$$\begin{aligned} \mathbb{P}\{\lambda \in \Lambda_{N(2^{-q})} \setminus \Psi_q\} &\leq 2^{3-q} 2^{\sqrt{N(2^{-q})}} e^{(\eta-\gamma)\sqrt{N(2^{-q})}} + \\ &+ 2^{2-q} 2^{N(2^{-q})+1} e^{(\eta-\gamma)(N(2^{-q})+1)} \leq \frac{O(1)}{q^2} \end{aligned}$$

which is summable over q . The lemma is proved. \square

For a given ε , let $\lambda \in \Psi_{\log(\varepsilon^{-1})}$ and define

$$M_{\varepsilon, N(\varepsilon)}(\lambda) := \Psi_{\log(\varepsilon^{-1})} \cap \left\{ \begin{array}{l} |\lambda_1 - \tilde{\lambda}_1| < 2\varepsilon \\ \tilde{\lambda} : |\lambda_k - \tilde{\lambda}_k| < 2^{k+1} (e^{\eta-\gamma})^{\sqrt{N}} \varepsilon \quad \text{if } k = 2, \dots, \lfloor \sqrt{N} \rfloor - 1 \\ |\lambda_k - \tilde{\lambda}_k| < \frac{2}{e^{\eta-\gamma}} (2 e^{\eta-\gamma})^k \varepsilon \quad \text{if } k = \lfloor \sqrt{N} \rfloor, \dots, N \end{array} \right\} \quad (29)$$

We now show that if $\tilde{\lambda} \notin M(\lambda)$ then $d_{\text{H}}(\mathcal{C}_\lambda, \mathcal{C}_{\tilde{\lambda}}) \geq \varepsilon$. This follows from Lemma 4.1 and we now check the hypothesis to be satisfied.

If $\tilde{\lambda} \in \Psi_{\log(\varepsilon^{-1})} \setminus M(\lambda)$ then one of the conditions in (29) is violated. Following the notation of Lemma 4.1, we start with $I = I' = [0, 1]$. We take $H = [\frac{\lambda_1}{2}, 1 - \frac{\lambda_1}{2}]$ and $H' = [\frac{\tilde{\lambda}_1}{2}, 1 - \frac{\tilde{\lambda}_1}{2}]$. Then $|H| > 2\varepsilon$ and $|H'| > 2\varepsilon$ since λ and $\tilde{\lambda}$ are in $\Psi_{\log(\varepsilon^{-1})}$. If $|\lambda_1 - \tilde{\lambda}_1| \geq 2\varepsilon$, then $|\frac{\lambda_1}{2} - \frac{\tilde{\lambda}_1}{2}| \geq \varepsilon$, and Lemma 4.1 applies with $F = \mathcal{C}_\lambda$ and $F' = \mathcal{C}_{\tilde{\lambda}}$, implying $d_{\text{H}}(\mathcal{C}_\lambda, \mathcal{C}_{\tilde{\lambda}}) > \varepsilon$.

Assume that, for some $k = 2, \dots, \lfloor \sqrt{N} \rfloor - 1$, all conditions in (29) are satisfied up to $k-1$ and condition k is violated. Either there is an $\ell < k$ such that

$$\left| \frac{1}{2^\ell} \prod_{h=1}^{\ell} \lambda_h - \frac{1}{2^\ell} \prod_{h=1}^{\ell} \tilde{\lambda}_h \right| > \varepsilon,$$

in which case we define \hat{k} to be the smallest such ℓ . Or, if no such ℓ exists,

$$\begin{aligned} \left| \frac{1}{2^k} \prod_{h=1}^k \lambda_h - \frac{1}{2^k} \prod_{h=1}^k \tilde{\lambda}_h \right| &\geq \frac{|\lambda_k - \tilde{\lambda}_k|}{2^k} \prod_{h=1}^{k-1} \lambda_h - \frac{\tilde{\lambda}_k}{2^k} \left| \prod_{h=1}^{k-1} \lambda_h - \prod_{h=1}^{k-1} \tilde{\lambda}_h \right| \geq \\ &\geq 2 \varepsilon (e^{\eta-\gamma})^{\sqrt{N}} \prod_{h=1}^{k-1} \lambda_h - \varepsilon > \varepsilon \end{aligned}$$

where we have used (26) and the fact that the leftmost positive points up to the $(k-1)$ -th step of the construction are ε -close, and we set $\hat{k} = k$.

We will apply Lemma 4.1 with $I = J_1^{\hat{k}-1}(\lambda)$ and $I' = J_1^{\hat{k}-1}(\tilde{\lambda})$. We take

$$H = \left[\frac{1}{2^{\hat{k}}} \prod_{h=1}^{\hat{k}} \lambda_h, \left(1 - \frac{\lambda_{\hat{k}}}{2}\right) \frac{1}{2^{\hat{k}-1}} \prod_{h=1}^{\hat{k}-1} \lambda_h \right]$$

and

$$H' = \left[\frac{1}{2^{\hat{k}}} \prod_{h=1}^{\hat{k}} \tilde{\lambda}_h, \left(1 - \frac{\tilde{\lambda}_{\hat{k}}}{2}\right) \frac{1}{2^{\hat{k}-1}} \prod_{h=1}^{\hat{k}-1} \tilde{\lambda}_h \right].$$

We have $|H| > 2\varepsilon$ and $|H'| > 2\varepsilon$ since λ and $\tilde{\lambda}$ are in $\Psi_{\log(\varepsilon^{-1})}$, and we can apply Lemma 4.1 which gives $d_H(\mathcal{C}_\lambda, \mathcal{C}_{\tilde{\lambda}}) > \varepsilon$.

The same argument applies if the k -th condition with $k = \sqrt{N}, \dots, N$ is violated, and all conditions up to $k-1$ are satisfied. If the leftmost positive points up to the $(k-1)$ -th step of the construction are ε -close, we get

$$\left| \frac{1}{2^k} \prod_{h=1}^k \lambda_h - \frac{1}{2^k} \prod_{h=1}^k \tilde{\lambda}_h \right| \geq 2\varepsilon (e^{\eta-\gamma})^{k-1} \prod_{h=1}^{k-1} \lambda_h - \varepsilon > \varepsilon$$

by definition (25) of $\Lambda_{N(\varepsilon)}$ and using, as above, that all previous leftmost positive points are ε -close to each other. Again this implies that $d_H(\mathcal{C}_\lambda, \mathcal{C}_{\tilde{\lambda}}) > \varepsilon$.

Let us now estimate the measure of the set $M_{\varepsilon, N(\varepsilon)}(\lambda)$. By an easy computation based on the independence of the random variables $(\tilde{\lambda}_k)$, we obtain that for all $\lambda \in \Psi_{\log(\varepsilon^{-1})}$

$$\begin{aligned} &\mathbb{P}(M_{\varepsilon, N(\varepsilon)}(\lambda)) \\ &\leq \left(\max_{x \in [0,1]} f(x) \right)^{N(\varepsilon)} (2\varepsilon)^{N(\varepsilon)} (e^{\eta-\gamma})^{-\sqrt{N(\varepsilon)}} 2^{\sum_{k=2}^{N(\varepsilon)} k} (e^{\eta-\gamma})^{\sum_{k=1}^{\lfloor \sqrt{N(\varepsilon)} \rfloor} k} \\ &\leq (2 e^{\eta-\gamma})^{-\frac{N(\varepsilon)^2}{2} + O(N(\varepsilon))} = e^{-O(1)(\log(\varepsilon^{-1}))^2}, \end{aligned} \quad (30)$$

where we have used the definition (27) of $N(\varepsilon)$. Note that this estimate is uniform in $\lambda \in \Psi_{\log(\varepsilon^{-1})}$.

For a fixed $\delta \in (0, 1)$, define $a_q := q^{2-\delta}$. For any q , there are at most 2^{a_q} different binary programs $(\mathbf{P}_j)_{1 \leq j \leq 2^{a_q}}$ of length $a_q - 1$, which generate at most 2^{a_q} different sets $\mathbf{C}_j := \mathbf{C}(\mathbf{P}_j)$. We define

$$V_{q,j} := \{\lambda : d_{\mathbb{H}}(\mathcal{C}_\lambda, \mathbf{C}_j) < 2^{-q}\}.$$

Then

$$\{\lambda : \Delta(\mathcal{C}_\lambda, 2^{-q}) < a_q\} \subset \bigcup_{j=1}^{2^{a_q}} V_{q,j}.$$

We can write

$$\mathbb{P}\left(\bigcup_{j=1}^{2^{a_q}} V_{q,j}\right) \leq \mathbb{P}(\Lambda_{N(2^{-q})}^c) + \mathbb{P}(\Psi_q^c \cap \Lambda_{N(2^{-q})}) + \sum_{j=1}^{2^{a_q}} \mathbb{P}(V_{q,j} \cap \Psi_q).$$

Moreover if $V_{q,j} \cap \Psi_q \neq \emptyset$, there is a $\lambda \in \Psi_q$ such that $V_{q,j} \cap \Psi_q \subset M_{2^{-q}, N(2^{-q})}(\lambda)$. By Lemmas 4.3 and 4.4, and by (30) it follows that

$$\sum_q \sum_{j=1}^{2^{a_q}} \mathbb{P}(V_{q,j}) < \infty.$$

The result follows from Lemma 4.2.

4.4 Proof of Theorem 3.4

We first prove (12). Let ε be fixed. We show how to approximate the set \mathcal{C} within Hausdorff distance ε . We will give the proof for integer $k \geq 1$. The proof easily extends to functions whose k -th derivative is Hölder.

We can write the Taylor expansions of the maps ϕ_i at a point x_0

$$\phi_i(x) = \sum_{p=0}^{k-1} c_{i,p}(x_0) (x - x_0)^p + R_i(x, x_0) \quad \forall x \in [0, 1].$$

Moreover there exists a constant $K > 0$ such that $|R_i(x, x_0)| \leq K|x - x_0|^k$ for $x \in [0, 1]$, for all $i \in \mathcal{I}$ and $x_0 \in [0, 1]$.

Let $\varepsilon' = \varepsilon \frac{1-\rho}{M}$ for a constant M to be specified later on. We now construct a sequence of polynomials which approximate the maps $(\phi_i)_{i \in \mathcal{I}}$. If D is the box counting dimension of the Cantor set \mathcal{C} , we need for any $\delta > 0$ at most $N = O((\varepsilon')^{-\frac{D}{k}-\delta})$ intervals $(I_s)_{s=1, \dots, N}$ of size $(\varepsilon')^{\frac{1}{k}}$ to cover \mathcal{C} . Hence we can consider the maps $(\phi_i)_{i \in \mathcal{I}}$ restricted to the sets (I_s) . If y_s denotes the middle point of the interval I_s , let \tilde{y}_s be the approximation of the point y_s within a distance ε' . Then we define

$$\tilde{\phi}_i^s(x) = \sum_{p=0}^k \tilde{c}_{i,p}(y_s) (x - \tilde{y}_s)^p \quad \forall x \in [0, 1]$$

such that

$$|c_{i,p}(y_s) - \tilde{c}_{i,p}(y_s)| < \varepsilon' \quad \forall i \in \mathcal{I} \quad \forall p = 0, \dots, k \quad (31)$$

and they are contractions on \mathbb{R} with the same uniform contraction rate $\tilde{\rho} < 1$.

To construct an approximation of \mathcal{C} , we work on the boundary points of the intervals $J_{\omega_1^n}$ which all are in \mathcal{C} . Let us denote $J_{\omega_1^n} = [y_{\omega_1^n}^1, y_{\omega_1^n}^2]$. Since for any $n \in \mathbb{N}$ and any $\omega_1^n \in \mathcal{I}^n$ we have $y_{\omega_1^n}^\eta \in \mathcal{C}$ for $\eta = 1, 2$, we can associate to a given $y_{\omega_1^n}^\eta$ a sequence $\sigma_0^{n-1} \in \{1, \dots, N\}^{n-1}$ which specifies to which intervals of the cover $(I_s)_{s=1, \dots, N}$ the pre-images $y_{\omega_1^{n-1}}^\eta, y_{\omega_1^{n-2}}^\eta, \dots, y_{\omega_1}^\eta, y_{\#}^\eta$ of $y_{\omega_1^n}^\eta$ belong, where $y_{\#}^\eta \in \{0, 1\}$.

We now establish the analogue of (15) for the boundary points. Let us define

$$\tilde{y}_{\omega_1^n}^\eta := \tilde{\phi}_{\omega_n^{n-1}}^{\sigma_{n-1}} \circ \tilde{\phi}_{\omega_{n-1}^{n-2}}^{\sigma_{n-2}} \circ \dots \circ \tilde{\phi}_{\omega_1}^{\sigma_0}(y_{\#}^\eta)$$

then for all $n \in \mathbb{N}$ it holds

$$|y_{\omega_1^n}^\eta - \tilde{y}_{\omega_1^n}^\eta| < \left(k + 1 + K + \max_{i \in \mathcal{I}, s=1, \dots, N} \sum_{p=0}^k |c_{i,p}(y_s)| \right) \frac{\varepsilon'}{1 - \rho} \quad \forall \omega_1^n \in \mathcal{I}^n. \quad (32)$$

The proof is by induction. The first step ($n = 1$) follows by definition of the approximating polynomials, estimates (31) and properties of the remainder $R_i(x, x_0)$. This yields

$$\begin{aligned} |y_i^\eta - \tilde{y}_i^\eta| &= |\phi_i(y_{\#}^\eta) - \tilde{\phi}_i^{\sigma_0}(y_{\#}^\eta)| \\ &\leq \varepsilon' \sum_{p=0}^k (\varepsilon')^{\frac{p}{k}} + \sum_{p=0}^k |c_{i,p}(y_{\sigma_0})| |(y_{\#}^\eta - y_{\sigma_0})^p - (y_{\#}^\eta - \tilde{y}_{\sigma_0})^p| + K\varepsilon' \\ &\leq \left(k + 1 + K + \max_{i \in \mathcal{I}, s=1, \dots, N} \sum_{p=0}^k |c_{i,p}(y_s)| \right) \varepsilon'. \end{aligned}$$

The inductive step follows by using the triangle inequality

$$|y_{\omega_1^n}^\eta - \tilde{y}_{\omega_1^n}^\eta| \leq |\phi_{\omega_n}(y_{\omega_1^{n-1}}^\eta) - \tilde{\phi}_{\omega_n}^{\sigma_{n-1}}(y_{\omega_1^{n-1}}^\eta)| + |\tilde{\phi}_{\omega_n}^{\sigma_{n-1}}(y_{\omega_1^{n-1}}^\eta) - \tilde{\phi}_{\omega_n}^{\sigma_{n-1}}(\tilde{y}_{\omega_1^{n-1}}^\eta)|$$

together with

$$|\phi_{\omega_n}(y_{\omega_1^{n-1}}^\eta) - \tilde{\phi}_{\omega_n}^{\sigma_{n-1}}(y_{\omega_1^{n-1}}^\eta)| < \varepsilon' \left(k + 1 + K + \max_{i \in \mathcal{I}, s=1, \dots, N} \sum_{p=0}^k |c_{i,p}(y_s)| \right)$$

and

$$|\tilde{\phi}_{\omega_n}^{\sigma_{n-1}}(y_{\omega_1^{n-1}}^\eta) - \tilde{\phi}_{\omega_n}^{\sigma_{n-1}}(\tilde{y}_{\omega_1^{n-1}}^\eta)| < \rho |y_{\omega_1^{n-1}}^\eta - \tilde{y}_{\omega_1^{n-1}}^\eta|,$$

where ρ is the uniform contraction rate of the approximating polynomials.

Let us choose \bar{n} such that $\rho^{\bar{n}} < \frac{\varepsilon}{2}$ for all $n \geq \bar{n}$. Then we define the set

$$\mathbf{C} := \bigcup_{\omega_1^{\bar{n}} \in \mathcal{I}^{\bar{n}}} \left(\tilde{y}_{\omega_1^{\bar{n}}}^1 \cup \tilde{y}_{\omega_1^{\bar{n}}}^2 \right)$$

and we claim that $d_{\text{H}}(\mathcal{C}, \mathbf{C}) < \varepsilon$. Indeed, by definition of \bar{n} , any point in the Cantor set \mathcal{C} is at most at distance $\frac{\varepsilon}{2}$ from a point in the boundary of one of the sets $J_{\omega_1^{\bar{n}}}$. Moreover, by construction of the points $\tilde{y}_{\omega_1^{\bar{n}}}^{\eta}$ we have (32), hence the claim follows since

$$\begin{aligned} d_{\text{H}}(\mathcal{C}, \mathbf{C}) &\leq \sup_{x \in \mathcal{C}} \inf_{\omega_1^{\bar{n}} \in \mathcal{I}^{\bar{n}}} |x - \tilde{y}_{\omega_1^{\bar{n}}}^{\eta}| \\ &\leq \sup_{x \in \mathcal{C}} \inf_{\omega_1^{\bar{n}} \in \mathcal{I}^{\bar{n}}} \left(|x - y_{\omega_1^{\bar{n}}}^{\eta}| + |y_{\omega_1^{\bar{n}}}^{\eta} - \tilde{y}_{\omega_1^{\bar{n}}}^{\eta}| \right) < \varepsilon \end{aligned}$$

provided that we choose

$$M := 2 \left(k + 1 + K + \max_{i \in \mathcal{I}, s=1, \dots, N} \sum_{p=0}^k |c_{i,p}(y_s)| \right).$$

Let us define the program \mathbf{P} that contains the numbers ε , ρ , D , M , K and k , and such that it specifies all the necessary coefficients $\tilde{c}_{i,p}$, makes the computation to obtain N and the approximated points \tilde{y}_s , and moreover it makes the computations to obtain \bar{n} and the points $\tilde{y}_{\omega_1^{\bar{n}}}^{\eta}$. The binary length $\ell(\mathbf{P})$ satisfies

$$\ell(\mathbf{P}) \preccurlyeq \varepsilon^{-\frac{D}{k} - \delta}$$

since ε is specified with $O(\log(\varepsilon^{-1}))$ bits; ρ , D , M , K and k do not depend on ε and can be approximated by rational numbers. The coefficients $\tilde{c}_{i,p}$ and the points \tilde{y}_s are constructed as in the previous proofs with precision ε' , hence each of them needs $O(\log(\varepsilon^{-1}))$ bits of information and their number is $N(k+2) = O(\varepsilon^{-\frac{D}{k} - \delta})$. Finally all the computations to obtain \mathbf{C} need $O(1)$ bits of instructions. Hence (12) follows.

We now prove (13). We define a class of particular scaling functions $S(\alpha)$ on \mathcal{I}^{∞} to construct differentiable Cantor sets with the given complexity.

Let us denote by $\mathcal{I}^* := \cup_{n \in \mathbb{N}} \mathcal{I}^n$ the countable set of finite strings s written using the alphabet \mathcal{I} . Let $(\lambda_s)_{s \in \mathcal{I}^*}$ be a family of independent identically distributed random variables with values in the interval $(0, 1)$ and absolutely continuous distribution with density $f(x)$ bounded above and below away from zero. Note that the empty string $\#$ belongs to \mathcal{I}^* and therefore there is an associated random variable $\lambda_{\#}$.

Let $0 < \zeta < 1$, $0 < \rho < 1$ and $\rho < \theta < 1$ be given constants, with ρ determining the contraction rate. ζ will be chosen small enough later on.

We will only consider central Cantor sets, namely the two components of the scaling functions will be equal. We define the scaling function

$$S_\lambda(\alpha) := \rho + \zeta \sum_{q=1}^{\infty} \theta^{q-1} \lambda_{\alpha_1^q} \quad \forall \alpha \in \mathcal{I}^\infty$$

which depends on the realisation of the family (λ_s) . We remark that for any realisation it holds

$$\rho \leq S_\lambda(\alpha) \leq \rho + \frac{\zeta}{1-\theta} \quad \forall \alpha \in \mathcal{I}^\infty. \quad (33)$$

Hence if ζ is small enough, the rate of contraction is almost ρ . It is also useful to define the truncated scaling function \tilde{S}_λ by

$$\tilde{S}_\lambda(\omega_1^m) := \rho + \zeta \sum_{q=1}^m \theta^{q-1} \lambda_{\omega_1^q} \quad \forall m \geq 0. \quad (34)$$

Using the relations

$$\frac{|J_{i\omega_1^m}|}{|J_{\omega_1^m}|} = \tilde{S}_\lambda(\omega_1^m) \quad \forall i \in \mathcal{I} \quad (35)$$

we can construct a central Cantor set \mathcal{C}_λ generated by the scaling function S_λ . From (33) it follows that the Cantor set \mathcal{C}_λ has box counting dimension $D(\zeta)$ which satisfies

$$D(\zeta) = -\frac{\log 2}{\log \rho} + O(\zeta) \quad \text{as } \zeta \rightarrow 0. \quad (36)$$

We now consider the differentiability of the IHIFS generating \mathcal{C}_λ . By (33) and the definition (6) of d_S it follows

$$\rho^n \leq d_S(\omega, \tilde{\omega}) \leq \left(\rho + \frac{\zeta}{1-\theta} \right)^n \quad n = |\omega \cap \tilde{\omega}|$$

for any $\omega, \tilde{\omega} \in \mathcal{I}^\infty$. Moreover for any $\omega \neq \tilde{\omega}$ it holds

$$|S_\lambda(\omega) - S_\lambda(\tilde{\omega})| \leq 2\zeta \frac{\theta^{n+1}}{1-\theta} \quad n = |\omega \cap \tilde{\omega}|.$$

Hence for $m > n$ we have diffeomorphisms from $A(\omega_1^m)$ into $A(\tilde{\omega}_1^m)$ with derivatives bounded by a constant $C(\omega_1^m, \tilde{\omega}_1^m) = O(\theta^{n+1})$. These facts together with relation (7) imply that the Cantor set \mathcal{C}_λ is of class C^k with

$$k = 1 + \frac{\log \theta}{\log \rho} > 1. \quad (37)$$

Let $0 < \varepsilon < 1$ be fixed and small enough depending on the constants ρ, θ, ζ . Let $\lambda := (\lambda_s)$ and $\lambda' := (\lambda'_s)$ denote two different realisations of the family of random variables. We give a condition on λ and λ' to have $d_{\text{H}}(\mathcal{C}_\lambda, \mathcal{C}_{\lambda'}) > \varepsilon$. We denote

$$\bar{p} = \left\lceil \frac{\log(C\varepsilon \zeta^{-1})}{\log(\rho\theta)} \right\rceil \quad (38)$$

where C is a positive constant (independent of ε) to be specified later on.

For any $\sigma \in \mathcal{I}^*$, we denote by J_σ and J'_σ the intervals associated to σ in \mathcal{C}_λ and $\mathcal{C}_{\lambda'}$ respectively.

Lemma 4.5. *Assume there is $0 \leq p \leq \bar{p}$ satisfying*

$$\max_{\omega_1^p \in \mathcal{I}^p} d_{\text{H}}(J_{\omega_1^p}, J'_{\omega_1^p}) > \varepsilon. \quad (39)$$

Then $d_{\text{H}}(\mathcal{C}_\lambda, \mathcal{C}_{\lambda'}) > \varepsilon$.

Proof. Denote by p the smallest integer for which the above inequality holds and by $\omega_1^p \in \mathcal{I}^p$ the string realising the maximum. We apply Lemma 4.1 with $I = J_{\omega_1^{p-1}}$ and $I' = J'_{\omega_1^{p-1}}$. The hypotheses on I and I' follow by the fact that (39) is violated up to $p-1$. The gaps H and H' have size at least

$$\rho^{p-1} > (\rho\theta)^{\bar{p}} = C\varepsilon \zeta^{-1} > 2\varepsilon$$

for ε small enough if $C\zeta^{-1} > 2$.

Finally since $d_{\text{H}}(J_{\omega_1^p}, J'_{\omega_1^p}) > \varepsilon$ we have all the hypotheses of Lemma 4.1. Hence the lemma follows. \square

Lemma 4.6. *Assume that there exists $0 \leq p \leq \bar{p}$ and a sequence $\omega_1^p \in \mathcal{I}^p$ such that*

$$|\lambda_{\omega_1^p} - \lambda'_{\omega_1^p}| > (\rho\theta)^{-p} \frac{(4 + 2\rho^2)\varepsilon}{\zeta}.$$

Then $d_{\text{H}}(\mathcal{C}_\lambda, \mathcal{C}_{\lambda'}) > \varepsilon$.

Proof. Denote by p the smallest integer for which the above inequality holds. It is enough to assume that for any $q < p$ condition (39) is not verified, otherwise the proof follows immediately from Lemma 4.5.

Let x (respectively x') be the point in the boundary of $J_{\omega_1^p}$ (respectively $J'_{\omega_1^p}$) which is not in the boundary of $J_{\omega_2^p}$ (respectively $J'_{\omega_2^p}$). Let y (respectively y') be the other boundary point of $J_{\omega_2^p}$ (respectively $J'_{\omega_2^p}$). Since by the recursive assumption $|y - y'| \leq \varepsilon$, we have

$$|x - x'| \geq |x - y - (x' - y')| - \varepsilon.$$

Now since $|x - y| = |J_{\omega_1^p}|$ and $(x - y)$ and $(x' - y')$ have the same sign, we get

$$|x - x'| \geq \left| |J_{\omega_1^p}| - |J'_{\omega_1^p}| \right| - \varepsilon$$

This can also be written

$$\begin{aligned} |x - x'| &\geq \left| \tilde{S}_\lambda(\omega_2^p) |J_{\omega_2^p}| - \tilde{S}_{\lambda'}(\omega_2^p) |J'_{\omega_2^p}| \right| - \varepsilon \geq \\ &\geq \left| \tilde{S}_\lambda(\omega_2^p) - \tilde{S}_{\lambda'}(\omega_2^p) \right| \frac{|J_{\omega_2^p}| + |J'_{\omega_2^p}|}{2} - \frac{\tilde{S}_\lambda(\omega_2^p) + \tilde{S}_{\lambda'}(\omega_2^p)}{2} \left| |J_{\omega_2^p}| - |J'_{\omega_2^p}| \right| - \varepsilon \geq \\ &\geq \left| \tilde{S}_\lambda(\omega_2^p) - \tilde{S}_{\lambda'}(\omega_2^p) \right| \frac{|J_{\omega_2^p}| + |J'_{\omega_2^p}|}{2} - 3\varepsilon \geq \rho^p \left| \tilde{S}_\lambda(\omega_2^p) - \tilde{S}_{\lambda'}(\omega_2^p) \right| - 3\varepsilon \end{aligned}$$

since $d_H(J_{\omega_2^p}, J'_{\omega_2^p}) \leq \varepsilon$ and $S_\lambda \geq \rho$. From (34) we get

$$\tilde{S}_\lambda(\omega_2^p) - \tilde{S}_{\lambda'}(\omega_2^p) = \zeta \sum_{q=2}^p \theta^{q-2} (\lambda_{\omega_2^q} - \lambda'_{\omega_2^q})$$

and from (35) we have

$$\begin{aligned} \zeta \sum_{q=2}^{p-1} \theta^{q-2} (\lambda_{\omega_2^q} - \lambda'_{\omega_2^q}) &= \frac{(|J_{\omega_1^{p-1}}| - |J'_{\omega_1^{p-1}}|)(|J_{\omega_2^{p-1}}| + |J'_{\omega_2^{p-1}}|)}{2 |J_{\omega_2^{p-1}}| |J'_{\omega_2^{p-1}}|} \\ &\quad + \frac{(|J_{\omega_1^{p-1}}| + |J'_{\omega_1^{p-1}}|)(|J'_{\omega_2^{p-1}}| - |J_{\omega_2^{p-1}}|)}{2 |J_{\omega_2^{p-1}}| |J'_{\omega_2^{p-1}}|}. \end{aligned}$$

Since $|J_{\omega_1^{p-1}}| \leq |J_{\omega_2^{p-1}}| \leq (\rho + \frac{\zeta}{1-\theta})^{p-2}$ we get (using again $d_H(J_{\omega_2^p}, J'_{\omega_2^p}) \leq \varepsilon$, and $S_\lambda \geq \rho$)

$$\left| \zeta \sum_{q=2}^{p-1} \theta^{q-2} (\lambda_{\omega_2^q} - \lambda'_{\omega_2^q}) \right| \leq 2 \rho^{-p+2} \varepsilon.$$

Hence

$$\left| \tilde{S}_\lambda(\omega_2^p) - \tilde{S}_{\lambda'}(\omega_2^p) \right| \geq \zeta \theta^{p-2} |\lambda_{\omega_1^p} - \lambda'_{\omega_1^p}| - 2 \rho^{-p+2} \varepsilon.$$

We conclude that

$$|x - x'| \geq \zeta (\rho\theta)^p |\lambda_{\omega_1^p} - \lambda'_{\omega_1^p}| - (3 + 2\rho^2)\varepsilon.$$

Therefore the lemma follows by applying Lemma 4.1. \square

We now want to estimate for a given realisation λ of the family (λ_s) the probability of the event

$$\mathcal{E}_{\bar{p}}(\lambda) = \sup_{0 \leq q \leq \bar{p}} (\rho\theta)^q \sup_{\sigma \in \mathcal{I}^q} |\lambda_\sigma - \lambda'_\sigma| \leq \frac{(4 + 2\rho^2) \varepsilon}{\zeta}.$$

By independence of the family (λ_s) we get

$$\mathbb{P}(\mathcal{E}_{\bar{p}}(\lambda)) \leq \prod_{q=0}^{\bar{p}} (C (\rho\theta)^{-q} \varepsilon \zeta^{-1})^{2^q}$$

where

$$C = (4 + 2\rho^2) \sup_{x \in [0,1]} f(x)$$

where f is the density of the random variables (λ_s) .

These relations imply that

$$\begin{aligned} \log \mathbb{P}(\mathcal{E}_{\bar{p}}(\lambda)) &\leq \sum_{q=0}^{\bar{p}} 2^q (\log(C \varepsilon \zeta^{-1}) - q \log(\rho\theta)) \\ &\leq 2^{\bar{p}+1} \log(\rho\theta) - (3 + \bar{p}) \log(\rho\theta). \end{aligned}$$

For a fixed $\delta \in (0, 1)$, define $a_\ell := 2^{\ell(-\delta + \frac{D}{k})}$ and choose $\varepsilon = 2^{-\ell}$. For any ℓ , there are at most 2^{a_ℓ} different binary programs $(P_j)_{1 \leq j \leq 2^{a_\ell}}$ of length $a_\ell - 1$, which generate at most 2^{a_ℓ} different sets $\mathbf{C}_j := \mathbf{C}(P_j)$. We define

$$V_{\ell,j} := \left\{ \lambda : d_{\mathbf{H}}(\mathcal{C}_\lambda, \mathbf{C}_j) < 2^{-\ell} \right\}.$$

Then

$$\left\{ \lambda : \Delta(\mathcal{C}_\lambda, 2^{-\ell}) < a_\ell \right\} \subset \bigcup_{j=1}^{2^{a_\ell}} V_{\ell,j}.$$

If $V_{\ell,j}$ is not empty, there exists λ such that $V_{\ell,j} \subset \mathcal{E}_{\bar{p}}(\lambda)$. Then

$$\begin{aligned} \mathbb{P} \left(\bigcup_{j=1}^{2^{a_\ell}} V_{\ell,j} \right) &\leq 2^{a_\ell} \mathbb{P}(\mathcal{E}_{\bar{p}}(\lambda)) \leq \\ &\leq O(1) \exp \left(2^{\ell(-\delta + \frac{D}{k})} \log 2 + 2^{\bar{p}+1} \log(\rho\theta) - (3 + \bar{p}) \log(\rho\theta) \right). \end{aligned}$$

Using (36), (37) and (38), it follows that for any $\delta > 0$ one can find $\zeta > 0$ small enough such that this is summable in ℓ . Hence we can apply Lemma 4.2 to complete the proof. \square

References

- [1] E. A. Asarin: *Complexity of uniform approximations of continuous functions* (Russian), Uspekhi Mat. Nauk. **39** (1984), 157–169.
- [2] M. F. Barnsley: “Fractals everywhere”. Second edition. Academic Press Professional, Boston, MA, 1993.

- [3] M. F. Barnsley, R. L. Devaney, B. Mandelbrot, H.-O. Peitgen, D. Saupe, R. Voss: “The science of fractal images”. Springer-Verlag, New York, 1988.
- [4] R. Bamón, C. G. Moreira, S. Plaza, J. Vera: *Differentiable structures of central Cantor sets*, Ergodic Theory & Dynam. Syst. **17** (1997), no. 5, 1027–1042.
- [5] T. Bedford, A. M. Fisher: *Ratio geometry, rigidity and the scenery process for hyperbolic Cantor sets*, Ergodic Theory & Dynam. Syst. **17** (1997), no. 3, 531–564.
- [6] George D. Birkhoff: “Aesthetic Measure”. Cambridge Massachusetts University Press, 1933.
- [7] M. Braverman: *Parabolic Julia sets are polynomial time computable*, Nonlinearity **19** (2006), no. 6, 1383–1401.
- [8] J.-Y. Cai, J. Hartmanis: *On Hausdorff and topological dimensions of the Kolmogorov complexity of the real line*, J. Comput. System Sci. **49** (1994), no. 3, 605–619.
- [9] G. J. Chaitin: *Algorithmic information theory*. IBM J. Res. Develop. **21** (1977), no. 4, 350–359.
- [10] A. Dembo, O. Zeitouni: “Large deviations techniques and applications”, Second edition. Applications of Mathematics **38**. Springer-Verlag, New York, 1998.
- [11] K. Falconer: “Fractal geometry”. Mathematical foundations and applications. Second edition. John Wiley & Sons, Inc., Hoboken, NJ, 2003. “Techniques in fractal geometry”. John Wiley & Sons, Ltd., Chichester, 1997.
- [12] S. Galatolo: *Pointwise information entropy for metric spaces*, Nonlinearity **12** (1999), no. 5, 1289–1298
- [13] A. N. Kolmogorov: *Three approaches to the quantitative definition of information*. Internat. J. Comput. Math. **2** (1968), 157–168.
- [14] A. N. Kolmogorov, V. T. Tihomirov: ε -entropy and ε -capacity of sets in function spaces, in “Selected works of A.N.Kolmogorov, Vol. III”, A.N. Shiriyayev Ed., Dordrecht, Kluwer, 1993.
- [15] M. Li, P. Vitányi: “An Introduction to Kolmogorov Complexity and Its Applications”, second edition, GTCS, Springer-Verlag, 1997.
- [16] J. H. Lutz, E. Mayordomo: *Dimensions of points in self-similar fractals*, SIAM J. Comput. **38** (2008), no. 3, 1080–1112

- [17] B. Mandelbrot: “The fractal geometry of nature”, W. H. Freeman and Co., San Francisco, Calif., 1982.
- [18] P. Mattila: “Geometry of sets and measures in Euclidean spaces, Fractals and rectifiability”. Cambridge Studies in Advanced Mathematics **44**. Cambridge University Press, Cambridge, 1995.
- [19] F. Przytycki, F. Tangerman: *Cantor sets in the line: scaling functions and the smoothness of the shift-map*, Nonlinearity **9** (1996), no. 2, 403–412.
- [20] D. Sullivan: *Differentiable structures on fractal-like sets, determined by intrinsic scaling functions on dual Cantor sets*, in “The mathematical heritage of Hermann Weyl” (Durham, NC, 1987), 15–23, Proc. Sympos. Pure Math., 48, Amer. Math. Soc., Providence, RI, 1988.
- [21] K. Weihrauch: “Computable Analysis”, Springer, Berlin, 2000.