Random subsets of Cantor sets generated by trees of coin flips

Pieter Allaart and Taylor Jones

Abstract. We introduce a natural way to construct a random subset of a homogeneous Cantor set C in $[0, 1]$ via random labelings of an infinite M-ary tree, where $M \geq 2$. The Cantor set C is the attractor of an equicontractive iterated function system $\{f_1, \ldots, f_N\}$ that satisfies the open set condition with $(0, 1)$ as the open set. For a fixed probability vector $(p_1, \ldots, p_N)$, each edge in the infinite M-ary tree is independently labeled i with probability p_i, for all $i = 1, 2, \ldots, N$. Thus, each infinite path in the tree receives a random label sequence of numbers from $\{1, 2, \ldots, N\}$. We define F to be the (random) set of those points $x \in C$ which have a coding that is equal to the label sequence of some infinite path starting at the root of the tree. The set F may be viewed as a statistically self-similar set with extreme overlaps, and as such, its Hausdorff and box-counting dimensions coincide. We prove non-trivial upper and lower bounds for this dimension, and obtain the exact dimension in a few special cases. For instance, when $M = N$ and $p_i = 1/N$ for each i, we show that F is almost surely of full Hausdorff dimension in C but of zero Hausdorff measure in its dimension. For the case of two maps and a binary tree, we also consider deterministic labelings of the tree where, for a fixed integer $m \geq 2$, every mth edge is labeled 1, and compute the exact Hausdorff dimension of the resulting subset of C.

1. Introduction

Consider an iterated function system (IFS) $\{f_1, \ldots, f_N\}$ consisting of N similar contractions $f_i : [0, 1] \to [0, 1]$, $i = 1, 2, \ldots, N$ such that $f_i((0, 1)) \cap f_j((0, 1)) = \emptyset$ for all $i \neq j$. In other words, $\{f_1, \ldots, f_N\}$ satisfies the open set condition with the open set being an interval. We assume furthermore that the IFS is equicontractive; that is, there is a constant $0 < r \leq 1/N$ such that

$$|f_i(x) - f_i(y)| = r|x - y| \quad \text{for all } x, y \in [0, 1], \ i = 1, 2, \ldots, N.$$

The attractor of the IFS $\{f_1, \ldots, f_N\}$ is a nonempty compact set C satisfying

$$C = \bigcup_{i=1}^{N} f_i(C).$$

When $r < 1/N$, C is a Cantor set with

$$\dim_H C = \dim_B C = -\frac{\log N}{\log r},$$

where $\dim_H$ and $\dim_B$ denote Hausdorff and box-counting dimensions, respectively. When $r = 1/N$, C is simply the unit interval $[0, 1]$.

We are interested in certain random subsets of C that are generated by trees of coin flips. The idea – made precise in the next section but outlined here for the special case $N = 2$ – is simple: In a full infinite binary tree, we randomly label each edge with either 1 with probability p or 2 with probability $1 - p$, where $p \in (0, 1)$ is a fixed parameter. Each path down the tree from the root then has a random infinite label sequence which is the coding of a point in C. We let F denote the random subset of C of all points obtained in this way. We show that when $p = 1/2$, F has almost surely the same Hausdorff dimension as C itself, but with zero Hausdorff measure in its dimension. In particular, when $r = 1/2$, our construction generates a random subset of the unit interval which has Lebesgue measure zero but full Hausdorff dimension one almost surely. When $p \neq 1/2$, we obtain instead a random subset of C with strictly smaller Hausdorff (and upper box-counting) dimension. For this biased case, we are, unfortunately, only able to obtain rough (but non-trivial) bounds on the dimension. These are given in Theorem 2.2 below, which we prove in Sections 3 and 4. In fact, we develop our bounds for arbitrary $N \geq 2$ and for random subsets obtained by random labelings of M-ary trees, where $M \geq 2$.

The random subset F may be viewed as a statistically self-similar set with extreme overlaps, as we explain in the next section. Statistically self-similar sets were introduced in the 1980s by Falconer [6], Graf [10] and Mauldin and Williams [17], and have since been studied by many other authors. In almost all of these papers, some separation condition, such as the open set condition, is assumed. While there is a significant body of literature on dimension computations for deterministic self-similar sets with overlaps, little appears to be known about the random case in the absence of a separation condition. With this article, we hope to make a modest contribution to this line of research, although we did not manage to compute the exact dimension of F.

We observe that our setup is essentially the same as that of Benjamini et al. [2]. However, their principal focus was on the overlapping case (i.e., $r > 1/N$), and instead of the random set F, they were mainly interested in the stochastically self-similar measure μ supported by F. This random measure μ is in fact precisely the measure m_ω that we use in the proof of the lower bound; see Section 3. Benjamini et al. do not give dimension estimates for $r < 1/N$, and indeed, our results here are nearly disjoint from theirs.

We consider, in Section 6, a deterministic version of our randomization scheme, corresponding to a sequence of coin flips in which every mth flip is heads, and compute exactly the Hausdorff dimension of the resulting subset of C. Interestingly, this dimension is not strictly monotone in m.

1.1. Models of random fractals

Following Hutchinson's result [11] on the Hausdorff dimension of self-similar fractals generated by an IFS satisfying the open set condition, there was a rapid rise in analogous results on a broad class of random constructions known as *statistically self-similar fractals*. Below, we discuss a few models from the literature for constructing a random Cantor set in the line. The randomness we are interested in resides in the process of deletion of intervals, as opposed to position or length of intervals (considered in [13, 14]).

Perhaps the best-known model of random Cantor sets begins with a partition of the unit interval into $M \geq 2$ disjoint subintervals of length $1/M$, and a vector of probabilities $\mathbf{p} = (p_1, \ldots, p_M) \in [0, 1]^M$ (not necessarily a probability vector). At stage 1, the ith interval in the partition is either kept or discarded with probabilities p_i and $1 - p_i$, respectively. Within each non-discarded subinterval this process is repeated proportionally, and independently of past stages. The limit set, conditioned on non-extinction, results in a random Cantor-like fractal F, say. In [6, 10, 17], it was determined independently that $\dim_H F = \frac{1}{M} \log\left(\sum_{i=1}^{M} p_i\right)$ almost surely on $F \neq \varnothing$. This model was also used in [5] to study the question of when the arithmetic difference of two random Cantor sets contains an interval.

Another model for constructing random Cantor sets, which removes the need for conditioning on non-extinction, is to fix a sequence of integers $N_1, N_2, \ldots$ for which $1 \leq N_k \leq M$ for all k. At each stage k in the construction, and within each interval remaining at stage $k - 1$, randomly and independently choose N_k subintervals of size M^{-k}, then discard the rest. This was used in [4], where the almost sure Hausdorff dimension was computed, as well as a close variation of it in [18] in connection to tube null sets.

The random Cantor sets we consider here also never go extinct, but our construction involves a complex form of interdependence which implies that the number of "basic intervals" remaining at level n is neither Markovian nor a martingale, even after appropriate scaling. As a result, traditional martingale approaches or branching arguments are not applicable. In fact, we could think of only one natural (random) mass distribution on our random Cantor subset, which led us to a lower bound that is (for the biased coin case), unfortunately, significantly below the best upper bound we could find.

We point out that our construction also differs from the V-variable fractals of [1], and their generalization in [12] to random code tree fractals.

2. Notation and main result

For definiteness, we assume that $f_i([0, 1])$ lies to the left of $f_j([0, 1])$ when $i < j$. Further, for ease of presentation, we also assume that each f_i is orientation preserving; in other words, there are constants $b_1, \ldots, b_N$ such that $f_i(x) = rx + b_i$, for $i = 1, \ldots, N$. However, this is not essential for our main results; see Remark 4.2 below.

Our probability space is $\Omega := \{1, \ldots, N\}^{\mathbb{N}}$ equipped with its Borel σ-algebra $\mathcal{B}$ and a probability measure $\mu_{\mathbf{p}}$ to be described below. We also write $\Omega_n := \{1, \ldots, N\}^n$ for $n \in \mathbb{N}$, and $\Omega^* := \cup_{n \in \mathbb{N}} \Omega_n$. For any finite sequence $\omega^* = (\omega_1, \ldots, \omega_n) \in \Omega^*$ with $|\omega^*| = n$, we define the basic open set, or cylinder,

$$[\omega^*] := \{\omega \in \Omega : \omega|_n = \omega^*\},$$

where $\omega|_n$ denotes the restriction of ω to the first n letters. Let $\mathcal{S} := \{[\omega^*] : \omega^* \in \Omega^*\}$.

Denote by $\mathcal{P}$ the set of all probability vectors $\mathbf{p} = (p_1, \ldots, p_N) \in [0, 1]^N$ such that $\sum_{i=1}^N p_i = 1$. For $\mathbf{p} = (p_1, \ldots, p_N) \in \mathcal{P}$, we define a premeasure $\mu_{\mathbf{p}}$ on $\mathcal{S}$ by

$$\mu_{\mathbf{p}}([\omega^*]) := \prod_{i=1}^N p_i^{\#\{k \leq n : \omega_k^* = i\}}, \qquad \text{for } \omega^* = (\omega_1^*, \ldots, \omega_n^*),$$

and observe that $\mu_{\mathbf{p}}$ extends uniquely to a probability measure on $(\Omega, \mathcal{B})$ which we again denote by $\mu_{\mathbf{p}}$.

Next, we consider a second integer parameter $M \geq 2$, and let

$$\mathcal{I}_n := \{1, \ldots, M\}^n \quad (n \in \mathbb{N}), \qquad \mathcal{I}^* := \bigcup_{n=0}^{\infty} \mathcal{I}_n, \qquad \mathcal{I} := \{1, \ldots, M\}^{\infty}.$$

Here ,we set $\mathcal{I}_0 := \{\epsilon\}$, where ϵ denotes the empty word. We define an order $<_*$ on $\mathcal{I}^*$ in a natural way: For any pair of words $\mathbf{i} = (i_1, \ldots, i_m)$ and $\mathbf{j} = (j_1, \ldots, j_n)$ in $\mathcal{I}^*$,

(1) $\mathbf{i} <_* \mathbf{j}$ if $m < n$, and

(2) $\mathbf{i} <_* \mathbf{j}$ if $m = n$ and $\mathbf{i} <_{\text{LEX.}} \mathbf{j}$ in the lexicographic order.

We think of $<_*$ as ordering the M-ary tree from top to bottom, left to right.

Next, we define a function $\kappa : \mathcal{I}^* \to \mathbb{N}$ by

$$\kappa(\mathbf{i}) := 1 + \#\{\mathbf{j} \in \mathcal{I}^* : \mathbf{j} <_* \mathbf{i}\}, \qquad \mathbf{i} \in \mathcal{I}^*.$$

Note that κ is well defined as there are but finitely many words less than $\mathbf{i}$ in the $<_*$ order. For instance, when $M = 2$, the first few values of κ are $\kappa(1) = 1$, $\kappa(2) = 2$, $\kappa(1, 1) = 3$, $\kappa(1, 2) = 4$, $\kappa(2, 1) = 5$, etc. Observe that κ defines an order-preserving bijection from $(\mathcal{I}^*, <_*)$ to $(\mathbb{N}, <)$.

Now, we define the collection of random variables $X_{\mathbf{i}} : \Omega^* \to \{1, \ldots, N\}$ by

$$X_{\mathbf{i}}(\omega) = \omega_{\kappa(\mathbf{i})}, \qquad \mathbf{i} \in \mathcal{I}^*.$$

These are the random labels of the edges in the M-ary tree.

For any $\omega \in \Omega$, we define a subset of C by

$$F(\omega) := \left\{ x = \lim_{n \to \infty} f_{X_{\mathbf{i}|_1}(\omega)} \circ f_{X_{\mathbf{i}|_2}(\omega)} \circ \cdots \circ f_{X_{\mathbf{i}|_n}(\omega)}([0, 1]) : \mathbf{i} = (i_1, i_2, \ldots) \in \mathcal{I} \right\}.$$

(We emphasize that due to the randomization, different choices of the path $\mathbf{i}$ could yield the same point x in the limit set $F(\omega)$.)

2.1. F as the limit of a branching random walk

To help visualize the random subset $F(\omega)$, the reader may wish to imagine a kind of branching random walk. We illustrate this for the case $M = N = 2$. The walk begins with a single ancestor at time 0 occupying the interval $[0, 1]$. At each stage $n \geq 0$, each individual of generation n splits into two descendants of generation $n + 1$. If the individual occupies the interval I, its descendants, independently of each other, move either to the left subinterval $f_1(I)$ with probability p_1, or the right subinterval $f_2(I)$ with probability p_2. (Basically, this process is a "branching random walk with exponentially decreasing steps", as in [2].) Note that at stage n there are 2^n individuals, collectively occupying some (random) subcollection of the intervals $I_{\mathbf{i}}, \mathbf{i} \in \{1, 2\}^n$. Let E_n denote the union of the occupied intervals; then $F(\omega) = \bigcap_{n=1}^{\infty} E_n$.

The difficulty in computing the dimension of $F(\omega)$ stems from the fact that, in order to understand the probability law of the branching random walk below stage n, it is not sufficient to know merely the number of occupied intervals at stage n, but in fact, one needs to know the individual occupancy rates for each of the stage n intervals. As a result, we are able to obtain the exact dimension only for a few special cases.

2.2. F as a statistically self-similar set

We can also think of F as a statistically (or stochastically) self-similar set with extreme overlaps. The general setup of statistically self-similar sets, introduced by Graf in [10], is as follows: Let Sim denote the set of all similarities in $\mathbb{R}$, with the topology of uniform convergence, and let ν be a Borel probability measure on Sim^M. Consider

the space $\widetilde{\Omega} := (Sim^M)^{\mathcal{I}^*}$, with the product measure $\nu^{\mathcal{I}^*}$. For $\mathbf{i} \in \mathcal{I}^*$ and $\tilde{\omega} \in \widetilde{\Omega}$, put $S_{\mathbf{i},j}(\tilde{\omega}) := \tilde{\omega}(\mathbf{i})_j$, $j = 1, 2, \ldots, M$. When $\mathbf{i} = \epsilon$, we write simply S_j instead of $S_{\epsilon,j}$. Under $\nu^{\mathcal{I}^*}$, the random M-tuples of maps $(S_{\mathbf{i},1}, \ldots, S_{\mathbf{i},M})$, $\mathbf{i} \in \mathcal{I}^*$, are mutually independent and all have distribution ν. Now define the random mapping $\pi_{\tilde{\omega}} : \mathcal{I} \to \mathbb{R}$ by

$$\pi_{\tilde{\omega}}(\mathbf{i}) := \lim_{n \to \infty} S_{\mathbf{i}|_1}(\tilde{\omega}) \circ \cdots \circ S_{\mathbf{i}|_n}(\tilde{\omega})(0),$$

where $\mathbf{i}|_k$ denotes the prefix $i_1 \ldots i_k$ of the word $\mathbf{i} = i_1 i_2 \ldots$ Finally, set $K(\tilde{\omega}) := \pi_{\tilde{\omega}}(\mathcal{I})$. The set $K(\tilde{\omega})$ is statistically self-similar in the sense that

$$K(\tilde{\omega}) = \bigcup_{j=1}^{M} S_j(\omega)\big(K_j(\tilde{\omega})\big),$$

where $K_1, \ldots, K_M$ are independent copies of $K(\tilde{\omega})$ that are also independent of $(S_1, \ldots, S_M)$. If we take ν to be the product measure

$$\nu = \bigotimes_{i=1}^{M} (p_1 \delta_{f_1} + \cdots + p_N \delta_{f_N}), \tag{2.1}$$

then $K(\tilde{\omega})$ has the same probability distribution as $F(\omega)$. Note that for any $\tilde{\omega} \in \widetilde{\Omega}$ and any two words $\mathbf{i}, \mathbf{j} \in \mathcal{I}^*$ of equal length, either $S_{\mathbf{i}}(\tilde{\omega})(0,1) \cap S_{\mathbf{j}}(\tilde{\omega})(0,1) = \emptyset$ or $S_{\mathbf{i}}(\tilde{\omega})(0,1) = S_{\mathbf{j}}(\tilde{\omega})(0,1)$. Thus, $F(\omega)$ is a statistically self-similar set with "extreme overlaps".

It was shown by Falconer [7] (see also [8, Chapter 3]) that for any *deterministic* self-similar set, with or without overlaps, the Hausdorff and box-counting dimensions coincide. More recently, Liu and Wu [16] proved that the same is true for statistically self-similar sets, under a mild condition which is trivially satisfied by the measure ν in (2.1). Moreover, the dimension is almost surely constant. Therefore, we obtain the following result:

Proposition 2.1. *There is a constant $a \geq 0$ such that, $\mu_{\mathbf{p}}$-almost surely,* $\dim_H F = \underline{\dim}_B F = \overline{\dim}_B F = a$.

Note that for all ω, we have $F(\omega) \subseteq C$. On the other hand, F can always be covered by at most M^n intervals of length r^{-n} for each n. Thus, we have the trivial upper bound

$$\overline{\dim}_B F(\omega) \leq \min\left\{ -\frac{\log N}{\log r}, -\frac{\log M}{\log r} \right\}. \tag{2.2}$$

Our goal is to establish the following more precise dimension bounds:

Theorem 2.2.   (a) *Lower bound. We have, $\mu_{\mathbf{p}}$-almost surely,*

$$\dim_H F \geq \min\left\{ -\frac{\log M}{\log r}, \frac{\log(p_1^2 + \cdots + p_N^2)}{\log r} \right\}.$$

(b) *Upper bound. If*

$$\prod_{i=1}^{N} p_i^{-p_i} \leq M \leq \left(\prod_{i=1}^{N} p_i\right)^{-1/N}, \tag{2.3}$$

then there is a unique $\lambda \in [0, 1]$ *satisfying*

$$\sum_{i=1}^{N} p_i^{\lambda} \log(Mp_i) = 0, \tag{2.4}$$

and

$$\overline{\dim}_B F \leq -\frac{\lambda \log M + \log\left(\sum_{i=1}^{N} p_i^{\lambda}\right)}{\log r}, \qquad \mu_{\mathbf{p}}\text{-almost surely.} \tag{2.5}$$

Remark 2.3. (a) The upper bound in (2.5) really is better than the trivial bound (2.2). To see this, set

$$\phi(x) := x \log M + \log\left(\sum_{i=1}^{N} p_i^x\right), \qquad x \in \mathbb{R},$$

and observe that $\phi(0) = \log N$, $\phi(1) = \log M$. The function

$$\phi'(x) = \log M + \frac{\sum p_i^x \log p_i}{\sum p_i^x}$$

has exactly one zero namely at $x = \lambda$. Furthermore,

$$\phi'(0) = \log M + \frac{1}{N}\sum \log p_i \leq 0$$

and

$$\phi'(1) = \log M + \sum p_i \log p_i \geq 0,$$

where the last two inequalities follow from (2.3). Thus, $\phi(\lambda) \leq \min\{\log M, \log N\}$.

(b) Note that the inequalities in (2.3) are always satisfied when $M = N$, since

$$\sum_{i=1}^{N} p_i \log p_i \geq -\log N = \log\left(\frac{1}{N}\sum_{i=1}^{N} p_i\right) \geq \frac{1}{N}\sum_{i=1}^{N} \log p_i,$$

by concavity of the logarithm. But the range of values of M satisfying (2.3) can be much larger. For instance, take $N = 3$ and $\mathbf{p} = (0.05, 0.2, 0.75)$. Then a direct computation yields that $\prod_{i=1}^{N} p_i^{-p_i} \approx 1.9886$ and $\left(\prod_{i=1}^{N} p_i\right)^{-1/N} \approx 5.1087$, so M could be 2, 3, 4 or 5.

Unfortunately, if M is outside the range (2.3), our method yields nothing better than the trivial upper bound (2.2). However, sometimes that bound turns out to be the exact dimension:

Corollary 2.4 (Large N, small M). *If*

$$M \leq \min\left\{ \prod_{i=1}^{N} p_i^{-p_i}, \frac{1}{p_1^2 + \cdots + p_N^2} \right\},$$

then, $\mu_{\mathbf{p}}$-almost surely,

$$\dim_H F = \dim_B F = -\frac{\log M}{\log r}.$$

For example, the conclusion of Corollary 2.4 holds when $N = 3$, $M = 2$ and $\mathbf{p} = (0.2, 0.2, 0.6)$.

We observe that the inequality $M \leq (p_1^2 + \cdots + p_N^2)^{-1}$ can be satisfied only if either $M < N$, or $M = N$ and $p_i = 1/N$ for each i, since it is easy to see that $\sum_{i=1}^{N} p_i^2 \geq 1/N$, with equality if and only if $p_i = 1/N$ for all i.

When $N = 2$, the equation (2.4) can be solved explicitly, and we get

Corollary 2.5. *Take $N = 2$, and put $p := p_1$. If $p(1 - p) \leq M^{-2}$, then, $\mu_{\mathbf{p}}$-almost surely,*

$$\overline{\dim}_B F \leq \frac{\xi \log \xi + (1 - \xi) \log(1 - \xi)}{\log r},$$

where

$$\xi := \xi(p) := \frac{\log Mp}{\log p - \log(1 - p)}.$$

(When $M = 2$ and $p = 1/2$, we set $\xi(p) := 1/2$.)

Observe that the condition $p(1 - p) \leq M^{-2}$ is always satisfied when $M = 2$. We illustrate the upper and lower bounds for the case $N = M = 2$ in Figure 1. The bounds are plotted as a function of $p := p_1$. Observe that the upper bound tends to 0 as p tends to 0 or 1; however, it is still quite large even for values of p very close to 0 and 1. This stems from the fact that the upper bound is the composition of two functions which both have vertical tangent lines at 0 and 1.

Corollary 2.6. *When $M \geq N$ and $p_i = 1/N$ for each i, we have*

$$\dim_H F = \dim_B F = \dim_H C = -\frac{\log N}{\log r} \qquad \mu_{\mathbf{p}}\text{-almost surely.}$$

Remark 2.7. When $N = M$, the random subset F may also be viewed as an orthogonal projection of a statistically self-similar set in the plane satisfying the open set
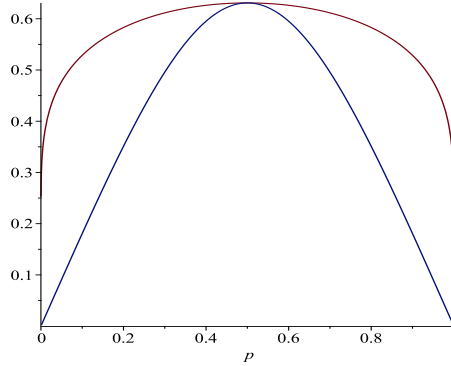
Figure 1. The upper and lower bounds of Theorem 2.2, with $N = M = 2$.

condition. We thank an anonymous referee for pointing this out. Define the maps

$$g_{i,j}(x, y) := \big(f_i(x), f_j(y) \big), \qquad i, j \in \{1, 2, \dots, M\}.$$

These maps take the unit square onto M^2 non-overlapping subsquares of width r, arranged in M rows and M columns. Now consider the setup from Subsection 2.2, but with maps on $\mathbb{R}^2$ instead of $\mathbb{R}$. We take the probability measure ν to be

$$\nu = \sum_{i_1, \dots, i_M} p_{i_1} \cdots p_{i_M} \delta_{(g_{i_1, 1}, g_{i_2, 2}, \dots, g_{i_M, M})}.$$

In essence, the measure ν picks out precisely one square in each row. Let $K(\tilde{\omega})$ be the resulting statistically self-similar set. It follows from the results in [6, 10, 17] that K has almost sure Hausdorff and box-counting dimension $-\log M / \log r$. The projection of K onto the y-axis is the full Cantor set C, while its projection onto the x-axis has the same probability distribution as our random subset F. Note that Marstrand's projection theorem implies that the projection of K in the "typical" direction (in the sense of Lebesgue measure) has Hausdorff dimension $-\log M / \log r = \dim_H C$. As a result, the horizontal direction in this context is "typical", whereas the vertical direction is not, except when $p_i = 1/M$ for each i.

3. Proof of the lower bound

Recall that every basic interval at level n in the construction of the deterministic Cantor set C is given by $I_{\mathbf{i}} := f_{i_1} \circ \cdots \circ f_{i_n}([0, 1])$ for some $\mathbf{i} \in \Omega_n$. We first define a random probability measure on $F(\omega)$:

For any $\mathbf{i} = (i_1, \ldots, i_n) \in \Omega_n$ and $\mathbf{j} = (j_1, \ldots, j_n) \in \mathcal{I}_n$, define the event

$$A(\mathbf{i}; \mathbf{j}) := \{\omega \in \Omega : (X_{j_1}(\omega), X_{j_1, j_2}(\omega), \ldots, X_{j_1, j_2, \ldots, j_n}(\omega)) = \mathbf{i}\}. \tag{3.1}$$

Then for $\mathbf{i} \in \Omega_n$, define

$$m_\omega(I_\mathbf{i}) := \frac{1}{M^n} \sum_{\mathbf{j} \in \mathcal{I}_n} 1_{A(\mathbf{i};\mathbf{j})}(\omega). \tag{3.2}$$

Intuitively, $m_\omega(I_\mathbf{i})$ is the relative number of branches labeled $\mathbf{i}$ in the tree at time n by the realization ω. It follows again from standard theorems that m_ω uniquely defines a measure on $F(\omega)$. We use this random measure along with the potential-theoretic method [9, Theorem 4.13] to prove the lower bound in Theorem 2.2.

Proof of Theorem 2.2, lower bound. Fix

$$t < s(\mathbf{p}) := \min\left\{-\frac{\log M}{\log r}, \frac{\log(p_1^2 + \cdots + p_N^2)}{\log r}\right\}. \tag{3.3}$$

We derive an estimate for the t-energy of m_ω, that is,

$$\Phi_t(m_\omega) := \iint\limits_{[0,1] \times [0,1]} |x - y|^{-t} dm_\omega(x) dm_\omega(y).$$

If x and y are two different points in $F(\omega)$, then there is a smallest basic interval (say $x \wedge y$) to which they both belong. If $x \wedge y = I_\mathbf{i}$ with $|\mathbf{i}| = n$, we define $k(x, y)$ to be the smallest integer k such that x and y are separated by at least one full basic interval of level $n + k$. Note that if $k(x, y) = k$, then

$$|x - y| \geq r^{n+k},$$

since each basic interval of level $n + k$ has length r^{n+k}. Without loss of generality, we may assume that $x < y$. If $k = 1$, then there are indices i_{n+1} and $i'_{n+1} \in \{1, \ldots, N\}$ such that $i'_{n+1} - i_{n+1} \geq 2$ and $x \in I_{\mathbf{i}, i_{n+1}}$, $y \in I_{\mathbf{i}, i'_{n+1}}$. If $k \geq 2$, there are two possibilities:

(i) $x \in I_{\mathbf{i}, i_{n+1}, N^{k-2}, i_{n+k}}$ and $y \in I_{\mathbf{i}, i_{n+1}+1, 1^{k-2}}$, where $i_{n+1} \in \{1, \ldots, N-1\}$ and $i_{n+k} \in \{1, \ldots, N-1\}$;

(ii) $x \in I_{\mathbf{i}, i_{n+1}, N^{k-2}}$ and $y \in I_{\mathbf{i}, i_{n+1}+1, 1^{k-2}, i'_{n+k}}$, where $i_{n+1} \in \{1, \ldots, N-1\}$ and $i'_{n+k} \in \{2, \ldots, N\}$.

Note that these two cases have some overlap; ignoring these overlaps leads at most to an overestimate of the energy.

We begin estimating the contribution of $\mathbf{i}$ to the t-energy of m_ω, $\Phi_t^{\mathbf{i}}(m_\omega)$, by

$$\Phi_t^{\mathbf{i}}(m_\omega) := \iint\limits_{\{(x,y):x\wedge y=I_{\mathbf{i}}\}} |x-y|^{-t} dm_\omega(x)dm_\omega(y) \tag{3.4}$$

$$= \sum_{k=1}^\infty \iint\limits_{\substack{x\wedge y=I_{\mathbf{i}}, \\ k(x,y)=k}} |x-y|^{-t} dm_\omega(x)dm_\omega(y)$$

$$\leq 2\sum_{k=2}^\infty r^{-(n+k)t}\left(\sum_{i_{n+1}=1}^{N-1}\sum_{i_{n+k}=1}^{N-1} m_\omega(I_{\mathbf{i},i_{n+1},N^{k-2},i_{n+k}})m_\omega(I_{\mathbf{i},i_{n+1}+1,1^{k-2}})\right. \tag{3.5}$$

$$+ \sum_{i_{n+1}=1}^{N-1}\sum_{i'_{n+k}=2}^{N} m_\omega(I_{\mathbf{i},i_{n+1},N^{k-2}})m_\omega(I_{\mathbf{i},i_{n+1}+1,1^{k-2},i'_{n+k}})\Bigg) \tag{3.6}$$

$$+ 2r^{-(n+1)t}\sum_{i_{n+1}=1}^{N-2}\sum_{i'_{n+1}=i_{n+1}+2}^{N} m_\omega(I_{\mathbf{i},i_{n+1}})m_\omega(I_{\mathbf{i},i'_{n+1}}). \tag{3.7}$$

We first estimate the expectation of the summands in (3.5). By the definition of m_ω, we have

$$m_\omega(I_{\mathbf{i},i_{n+1},N^{k-2},i_{n+k}}) = \frac{1}{M^{n+k}}\sum_{\mathbf{j}\in\mathcal{I}_{n+k}} 1_{A(\mathbf{i},i_{n+1},N^{k-2},i_{n+k};\mathbf{j})}(\omega),$$

$$m_\omega(I_{\mathbf{i},i_{n+1}+1,1^{k-2}} = \frac{1}{M^{n+k}}\sum_{\mathbf{j}'\in\mathcal{I}_{n+k-1}} 1_{A(\mathbf{i},i_{n+1}+1,1^{k-2};\mathbf{j}')}(\omega).$$

Thus, letting $E_{\mathbf{p}}$ denote the expectation operator associated with $\mu_{\mathbf{p}}$,

$$E_{\mathbf{p}}\big[m_\omega(I_{\mathbf{i},i_{n+1},N^{k-2},i_{n+k}})m_\omega(I_{\mathbf{i},i_{n+1}+1,1^{k-2}})\big]$$

$$= \frac{1}{M^{2n+2k-1}}\sum_{\mathbf{j}\in\mathcal{I}_{n+k}}\sum_{\mathbf{j}'\in\mathcal{I}_{n+k-1}} \mu_{\mathbf{p}}\big(A(\mathbf{i},i_{n+1},N^{k-2},i_{n+k};\mathbf{j})$$

$$\cap A(\mathbf{i},i_{n+1}+1,1^{k-2};\mathbf{j}'))$$

$$\leq \frac{1}{M^{2n+2k-1}} M^k p_{i_{n+1}} p_N^{k-2} p_{i_{n+k}} M^{k-1} p_{i_{n+1}+1} p_1^{k-2}\sum_{\mathbf{j},\mathbf{j}'\in\mathcal{I}_n} \mu_{\mathbf{p}}\big(A(\mathbf{i};\mathbf{j})\cap A(\mathbf{i};\mathbf{j}')\big)$$

$$\leq M^{-2n}(p_1 p_N)^{k-2} S_{\mathbf{p}}(\mathbf{i}),$$

where

$$S_{\mathbf{p}}(\mathbf{i}) := \sum_{\mathbf{j},\mathbf{j}'\in\mathcal{I}_n} \mu_{\mathbf{p}}\big(A(\mathbf{i};\mathbf{j})\cap A(\mathbf{i};\mathbf{j}')\big).$$

Similarly, for the summands in (3.6),

$$E_{\mathbf{p}}\big[m_\omega(I_{\mathbf{i},i_{n+1},N^{k-2}})m_\omega(I_{\mathbf{i},i_{n+1}+1,1^{k-2},i'_{n+k}})\big] \le M^{-2n}(p_1 p_N)^{k-2} S_{\mathbf{p}}(\mathbf{i}),$$

and for the summands in (3.7),

$$E_{\mathbf{p}}\big[m_\omega(I_{\mathbf{i},i_{n+1}})m_\omega(I_{\mathbf{i},i'_{n+1}})\big] \le M^{-2n} S_{\mathbf{p}}(\mathbf{i}).$$

Combining these estimates and noting that the double sums in (3.5)–(3.7) have at most N^2 terms, we obtain

$$E_{\mathbf{p}}\big[\Phi^{\mathbf{i}}_t(m_\omega)\big]$$

$$\le 4N^2 \sum_{k=2}^{\infty} r^{-(n+k)t} \cdot M^{-2n}(p_1 p_N)^{k-2} S_{\mathbf{p}}(\mathbf{i}) + 2N^2 r^{-(n+1)t} M^{-2n} S_{\mathbf{p}}(\mathbf{i})$$

$$\le 4N^2 M^{-2n} S_{\mathbf{p}}(\mathbf{i})\bigg(r^{-(n+2)t} \sum_{k=2}^{\infty} (r^{-t} p_1 p_N)^{k-2} + r^{-(n+1)t}\bigg)$$

$$= K(r^t M^2)^{-n} S_{\mathbf{p}}(\mathbf{i}) \tag{3.8}$$

for a certain constant K that depends on r and t, but not on $\mathbf{i}$. Note that the last summation in (3.8) converges because (3.3) implies

$$r^{-t} p_1 p_N < \frac{p_1 p_N}{p_1^2 + \cdots + p_N^2} \le \frac{p_1 p_N}{p_1^2 + p_N^2} \le \frac{1}{2} < 1.$$

To estimate the double sum $S_{\mathbf{p}}(\mathbf{i})$, we introduce the sets

$$W_n(q) := \big\{(\mathbf{j}, \mathbf{k}) \in \mathcal{I}_n \times \mathcal{I}_n : |\mathbf{j} \wedge \mathbf{k}| = q\big\}, \qquad q = 0, 1, \ldots, n,$$

where $\mathbf{j} \wedge \mathbf{k}$ is the longest common prefix of $\mathbf{j}$ and $\mathbf{k}$. Observe that for any $k \ge 1$, we have $\#W_k(0) = (M-1)M^{2k-1}$. Furthermore, we have the recursive relation $\#W_n(q) = M^q \cdot \#W_{n-q}(0)$ for $q < n$. Thus,

$$\#W_n(q) = M^q (M-1)M^{2n-2q-1} = (M-1)M^{2n-q-1}, \qquad q = 0, 1, \ldots, n-1.$$

For $q = n$, we have $\#W_n(n) = M^n$. In all cases, then, $\#W_n(q) \le M^{2n-q}$.

Before continuing with our previous calculation, we define

$$c_l(\mathbf{i}) := \#\{k \le n : i_k = l\}, \qquad \mathbf{i} \in \mathcal{I}_n, \quad n \in \mathbb{N}, \quad l = 1, \ldots, N. \tag{3.9}$$

We also write, for $\mathbf{i} \in \mathcal{I}_n$,

$$\mathbf{i}|^q := (i_{q+1}, \ldots, i_n),$$

to be the suffix of $\mathbf{i}$ after the qth digit (as opposed to the prefix $\mathbf{i}|_q$ up to q). Notice that $c_l(\mathbf{i}|_q) + c_l(\mathbf{i}|^q) = c_l(\mathbf{i})$. Now, we have

$$S_{\mathbf{p}}(\mathbf{i}) = \sum_{q=0}^{n} \sum_{(\mathbf{j},\mathbf{k}) \in W_n(q)} \mu_{\mathbf{p}}\big(A(\mathbf{i};\mathbf{j}) \cap A(\mathbf{i};\mathbf{k})\big). \tag{3.10}$$

Since the labels $(X_{j_1}, \ldots, X_{j_1,\ldots,j_n})$ and $(X_{k_1}, \ldots, X_{k_1,\ldots,k_n})$ of each pair of branches $(\mathbf{j}, \mathbf{k})$ in $W_n(q)$ agree up to index q, and are independent after q, we have

$$\mu_{\mathbf{p}}\big(A(\mathbf{i};\mathbf{j}) \cap A(\mathbf{i};\mathbf{k})\big) = \prod_{l=1}^{N} p_l^{c_l(\mathbf{i}|_q)} \cdot \prod_{l=1}^{N} p_l^{2c_l(\mathbf{i}|^q)},$$

for all $(\mathbf{j}, \mathbf{k}) \in W_n(q)$. Since there are at most M^{2n-q} pairs of words in $W_n(q)$, (3.10) gives

$$S_{\mathbf{p}}(\mathbf{i}) \leq \sum_{q=0}^{n} M^{2n-q} \prod_{l=1}^{N} p_l^{c_l(\mathbf{i}|_q)} \cdot \prod_{l=1}^{N} p_l^{2c_l(\mathbf{i}|^q)}.$$

Writing $\mathbf{i}' := \mathbf{i}|_q$ and $\mathbf{i}'' := \mathbf{i}|^q$, and noting that summing over all $\mathbf{i} \in \Omega_n$ is the same as summing independently over $\mathbf{i}' \in \Omega_q$ and $\mathbf{i}'' \in \Omega_{n-q}$, we obtain the estimate

$$\sum_{\mathbf{i} \in \Omega_n} S_{\mathbf{p}}(\mathbf{i}) \leq \sum_{q=0}^{n} M^{2n-q} \sum_{\mathbf{i}' \in \Omega_q} \prod_{l=1}^{N} p_l^{c_l(\mathbf{i}')} \cdot \sum_{\mathbf{i}'' \in \Omega_{n-q}} \prod_{l=1}^{N} p_l^{2c_l(\mathbf{i}'')}.$$

Now observe that

$$\sum_{\mathbf{i}' \in \Omega_q} \prod_{l=1}^{N} p_l^{c_l(\mathbf{i}')} = \sum_{(i_1,\ldots,i_q) \in \Omega_q} p_{i_1} \cdots p_{i_q} = (p_1 + \cdots + p_N)^q = 1$$

and similarly,

$$\sum_{\mathbf{i}'' \in \Omega_{n-q}} \prod_{l=1}^{N} p_l^{2c_l(\mathbf{i}'')} = (p_1^2 + \cdots + p_N^2)^{n-q}.$$

Hence,

$$\sum_{\mathbf{i} \in \Omega_n} S_{\mathbf{p}}(\mathbf{i}) \leq \sum_{q=0}^{n} M^{2n-q}(p_1^2 + \cdots + p_N^2)^{n-q} = M^n \sum_{j=0}^{n} \big[M(p_1^2 + \cdots + p_N^2)\big]^j.$$

Combining this with (3.4) and (3.8) yields

$$E_{\mathbf{p}}\big[\Phi_t(m_\omega)\big] = \sum_{n=0}^{\infty} \sum_{\mathbf{i} \in \Omega_n} E_{\mathbf{p}}\big[\Phi_t^{\mathbf{i}}(m_\omega)\big] \leq \sum_{n=0}^{\infty} K(r^t M^2)^{-n} \sum_{\mathbf{i} \in \Omega_n} S_{\mathbf{p}}(\mathbf{i})$$

$$\leq \sum_{n=0}^{\infty} K(r^t M)^{-n} \sum_{j=0}^{n} \big[M(p_1^2 + \cdots + p_N^2)\big]^j.$$

We must consider two cases:

Case 1. $M(p_1^2 + \cdots + p_N^2) \geq 1$. Then

$$E_{\mathbf{p}}\big[\Phi_t(m_\omega)\big] \leq \sum_{n=0}^{\infty} K(n+1)(r^t M)^{-n}\big[M(p_1^2 + \cdots + p_N^2)\big]^n$$

$$= \sum_{n=0}^{\infty} K(n+1)\big[r^{-t}(p_1^2 + \cdots + p_N^2)\big]^n < \infty,$$

because (3.3) implies $r^{-t} < (p_1^2 + \cdots + p_N^2)^{-1}$.

Case 2. $M(p_1^2 + \cdots + p_N^2) < 1$. Then $\sum_{j=0}^{n}[M(p_1^2 + \cdots + p_N^2)]^j$ is bounded by some constant $\tilde{K}$, and we get simply

$$E_{\mathbf{p}}\big[\Phi_t(m_\omega)\big] \leq K\tilde{K} \sum_{n=0}^{\infty}(r^{-t}M^{-1})^n < \infty,$$

since (3.3) implies $r^{-t} < M$.

Therefore, for any $t < s(\mathbf{p})$ we have that $E_{\mathbf{p}}[\Phi_t(m_\omega)] < \infty$, which means that $\Phi_t(m_\omega) < \infty$ for $\mu_{\mathbf{p}}$-almost every $\omega \in \Omega$. Hence, $\dim_H F(\omega) \geq s(\mathbf{p})$ almost surely. ∎

4. Proof of the upper bound

Proof of Theorem 2.2, *upper bound.* Assume the inequalities (2.3) hold. Uniqueness of the zero λ of (2.4) follows from a version of Descartes' rule of signs: Arrange the p_i's so that $p_1 \leq p_2 \leq \cdots \leq p_N$, and make the substitution $u := p_1^\lambda$. Then

$$\sum p_i^\lambda \log(Mp_i) = \sum u^{\beta_i} \log(Mp_i) =: g(u),$$

where $\beta_i := \log p_i / \log p_1$. Note that the exponents β_i decrease in i whereas the coefficients $\log(Mp_i)$ increase in i, so there is at most one sign change in the coefficients. Descartes' rule of signs does not require the exponents to be integers, so there is at most one positive solution to $g(u) = 0$.

For a word $\mathbf{i} = (i_1, \ldots, i_n) \in \Omega_n$, let $a_\mathbf{i}$ be the probability under $\mu_{\mathbf{p}}$ that at least one path of length n starting at the root of the M-ary tree has label sequence $\mathbf{i}$. That is,

$$a_\mathbf{i} := \mu_{\mathbf{p}}\bigg(\bigcup_{\mathbf{j} \in \mathcal{I}_n} A(\mathbf{i}; \mathbf{j})\bigg). \tag{4.1}$$

Since the probability of $A(\mathbf{i};\mathbf{j})$ is the same for each path $\mathbf{j}$, we have

$$a_{\mathbf{i}} \leq M^n \prod_{l=1}^{N} p_l^{c_l(\mathbf{i})}, \tag{4.2}$$

where $c_l(\mathbf{i})$ was defined in (3.9). Trivially, we also have $a_{\mathbf{i}} \leq 1$. Let $Z_n(\omega)$ denote the number of basic intervals at level n in the construction of C that intersect the limit set $F(\omega)$. Then

$$E_{\mathbf{p}}(Z_n) = \sum_{\mathbf{i} \in \Omega_n} a_{\mathbf{i}}. \tag{4.3}$$

Our goal is to bound the growth rate of $E_{\mathbf{p}}(Z_n)$. Note that

$$E_{\mathbf{p}}(Z_n) \leq \sum_{k_1 + \cdots + k_N = n} \binom{n}{k_1, \ldots, k_N} \min\left\{1, M^n \prod_{i=1}^{N} p_i^{k_i}\right\}, \tag{4.4}$$

by grouping together all label words of length n that have the same digit frequencies. This is a rough estimate because it does not take into account the overlaps among the events $A(\mathbf{i};\mathbf{j})$, $\mathbf{j} \in \mathcal{I}_n$. But understanding these overlaps appears to be a daunting task that we have not pursued further. We can, however, determine the exact exponential growth rate of the right-hand side of (4.4).

Since the number of summands in (4.4) is no greater than n^{N-1} (a polynomial in n), it suffices to find the largest summand and determine its growth rate in n. To begin, write $k_i = \alpha_i n$ for $i = 1, \ldots, N$. Then, we need to solve the optimization problem

$$\max\left\{\binom{n}{\alpha_1 n, \ldots, \alpha_N n} \min\left\{1, \left(M \prod_{i=1}^{N} p_i^{\alpha_i}\right)^n\right\} : \alpha_i \in [0, 1] \,\forall i, \sum_i \alpha_i = 1\right\}, \tag{4.5}$$

where, if $\alpha_i n$ is not an integer, we define $(\alpha_i n)!$ to be $\Gamma(\alpha_i n + 1)$, with Γ denoting the gamma function. Now observe that

$$M \prod_{i=1}^{N} p_i^{\alpha_i} = 1 \iff \sum_{i=1}^{N} \alpha_i \log p_i = -\log M,$$

and this last equation determines a hyperplane in $\mathbb{R}^N$. This hyperplane intersects the simplex

$$\Delta := \left\{(\alpha_1, \ldots, \alpha_N) \in [0, 1]^N : \sum_{i=1}^{N} \alpha_i = 1\right\},$$

and in fact, by (2.3), the point $\mathbf{p} = (p_1, \ldots, p_N)$ lies on the side of the hyperplane where $M \prod_{i=1}^{N} p_i^{\alpha_i} \geq 1$, and the point $(1/N, \ldots, 1/N)$ lies on the opposite side. The functions

$$\psi_1(\alpha_1, \ldots, \alpha_N) := \binom{n}{\alpha_1 n, \ldots, \alpha_N n}$$

and

$$\psi_2(\alpha_1, \ldots, \alpha_N) := \binom{n}{\alpha_1 n, \ldots, \alpha_N n} \prod_{i=1}^{N} p_i^{\alpha_i}$$

are both unimodal on Δ; ψ_1 is maximized at $(\alpha_1, \ldots, \alpha_N) = (1/N, \ldots, 1/N)$; and ψ_2 is maximized when $(\alpha_1 n, \ldots, \alpha_N n)$ is near the mode of the multinomial $(n; p_1, \ldots, p_N)$ distribution, i.e., when $(\alpha_1, \ldots, \alpha_N) \approx (p_1, \ldots, p_N)$. The error in this approximation goes to zero as $n \to \infty$. Since the global maximum of both ψ_1 and ψ_2 lies on the "wrong" side of the hyperplane $\sum_{i=1}^{N} \alpha_i \log p_i = -\log M$, the constrained maximum in (4.5) must be attained on this hyperplane. By Stirling's approximation,

$$\binom{n}{\alpha_1 n, \ldots, \alpha_N n} = Q_n(\alpha_1, \ldots, \alpha_n)\left(\alpha_1^{\alpha_1} \ldots \alpha_N^{\alpha_N}\right)^{-n},$$

where Q_n is a function such that both Q_n and $1/Q_n$ grow at most at a polynomial rate in n. Hence, to determine the exponential growth rate of the maximum in (4.5), we must solve the constrained optimization

$$\max\left\{-\sum_{i=1}^{N} \alpha_i \log \alpha_i : (\alpha_1, \ldots, \alpha_N) \in \Delta, \sum_{i=1}^{N} \alpha_i \log p_i = -\log M\right\}.$$

This is a straightforward Lagrange multiplier problem; we find that the maximum is attained when

$$\alpha_i = \frac{p_i^{\lambda}}{\sum_j p_j^{\lambda}}, \qquad i = 1, \ldots, N,$$

where λ satisfies the equation

$$\frac{\sum p_i^{\lambda} \log p_i}{\sum p_i^{\lambda}} = -\log M,$$

i.e., $\sum p_i^{\lambda} \log(M p_i) = 0$. A further calculation then yields that

$$-\sum \alpha_i \log \alpha_i = \lambda \log M + \log\left(\sum p_i^{\lambda}\right).$$

Denote this last expression by ρ. Then, we find that, for given $\varepsilon > 0$ and sufficiently large n, $E_{\mathbf{p}}(Z_n) \leq (\rho + \varepsilon)^n$. So

$$\mu_{\mathbf{p}}\left(Z_n > (\rho + 2\varepsilon)^n\right) \leq \frac{E_{\mathbf{p}}(Z_n)}{(\rho + 2\varepsilon)^n} \leq \frac{(\rho + \varepsilon)^n}{(\rho + 2\varepsilon)^n}.$$

Hence, the series $\sum_{n=1}^{\infty} \mu_{\mathbf{p}}(Z_n > (\rho + 2\varepsilon)^n)$ converges, so by the Borel–Cantelli lemma, $Z_n \leq (\rho + 2\varepsilon)^n$ for all sufficiently large n with probability one. Therefore,

$$\overline{\dim}_B F(\omega) \leq -\frac{\log(\rho + 2\varepsilon)}{\log r} \qquad \text{for } \mu_{\mathbf{p}}\text{-almost every } \omega,$$

and letting $\varepsilon \to 0$ along a discrete sequence gives

$$\overline{\dim}_B F(\omega) \leq -\frac{\log \rho}{\log r} \qquad \text{for } \mu_{\mathbf{p}}\text{-almost every } \omega,$$

as was to be shown. ∎

Remark 4.1. We briefly comment here on the difficulty of improving the upper bound. Consider the simplest case, $N = M = 2$, and set $p_1 = p$ and $p_2 = 1 - p$. We can then write down a pair of quadratic recursions for the probabilities $a_{\mathbf{i}}$: For $n \geq 2$,

$$\begin{aligned}
a_{1,i_2,\ldots,i_n} &= 2p(1-p)a_{i_2,\ldots,i_n} + p^2(2a_{i_2,\ldots,i_n} - a_{i_2,\ldots,i_n}^2) \\
&= 2pa_{i_2,\ldots,i_n} - p^2 a_{i_2,\ldots,i_n}^2,
\end{aligned} \tag{4.6}$$

and

$$\begin{aligned}
a_{2,i_2,\ldots,i_n} &= 2p(1-p)a_{i_2,\ldots,i_n} + (1-p)^2(2a_{i_2,\ldots,i_n} - a_{i_2,\ldots,i_n}^2) \\
&= 2(1-p)a_{i_2,\ldots,i_n} - (1-p)^2 a_{i_2,\ldots,i_n}^2.
\end{aligned} \tag{4.7}$$

For instance, (4.6) can be understood by looking at the three cases in Figure 2.
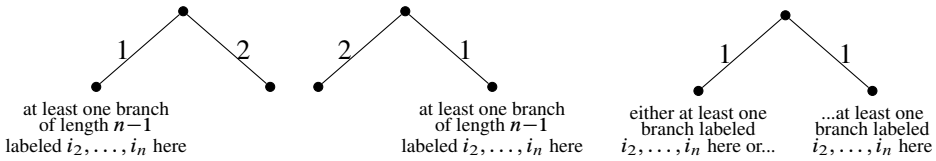
Figure 2. Three possibilities if a branch labeled $(1, i_2, \ldots, i_n)$ is in the tree at time n.

Ignoring the quadratic terms in the recursions would simply lead us back to (4.2). But a full understanding of their effects, averaged out over all paths $\mathbf{i}$, would entail a delicate analysis of the joint dynamics of the two quadratic maps $x \mapsto px(2 - px)$ and $x \mapsto (1 - p)x(2 - (1 - p)x)$, which appears to be extremely complicated. (For one thing, better bounds on $a_{\mathbf{i}}$ would depend not only on the number of 1's and 2's in $\mathbf{i}$, but also on the order in which the digits appear.)

Remark 4.2. While we assumed in Theorem 2.2 that the maps $f_1, \ldots, f_N$ are orientation preserving, the same result holds when one or more maps are instead made orientation-reversing. We briefly indicate here how the proof of the lower bound would need to be modified. For ease of presentation, consider again the case $N = M = 2$ and suppose $f_1(x) = -rx + b_1$ and $f_2(x) = rx + b_2$, where we still assume that $f_1([0, 1])$ lies to the left of $f_2([0, 1])$. In this case, (i) and (ii) at the beginning of the proof of the lower bound in Theorem 2.2 become, for $k \geq 3$:

(i) $x \in I_{\mathbf{i},1,1,2^{k-3},1}$ and $y \in I_{\mathbf{i},2,1,2^{k-3}}$;

(ii) $x \in I_{\mathbf{i},1,1,2^{k-3}}$ and $y \in I_{\mathbf{i},2,1,2^{k-3},1}$.

For $k = 2$, we have $x \in I_{\mathbf{i},1,2}$ and $y \in I_{\mathbf{i},2}$, or $x \in I_{\mathbf{i},1}$ and $y \in I_{\mathbf{i},2,2}$. Put $p := p_1$. The constant K from (3.8) is now replaced with the sum

$$\sum_{k=3}^{\infty} r^{-kt} p^4 (1-p)^{2(k-3)+1}$$

(plus an unimportant term for $k = 2$), which converges for

$$t < s(p) := \frac{\log\left(p^2 + (1-p)^2\right)}{\log r},$$

because the latter implies

$$r^{-t}(1-p)^2 < \frac{(1-p)^2}{p^2 + (1-p)^2} < 1.$$

Other combinations of orientations can be dealt with similarly. Clearly, the orientations of the maps have no bearing on the proof of the upper bound.

5. Hausdorff measure in the symmetric case

Throughout this section, we assume that we are in the symmetric case where all of the maps $f_1, \dots, f_N$ are chosen with the same probability; that is, $\mathbf{p} = (1/N, \dots, 1/N)$.

Corollary 2.6 states that if $M \geq N$, then the Hausdorff and box-counting dimensions of F are almost surely equal to $-\log N / \log r$, the dimension of C. In this section, we consider the Hausdorff *measure* of F in the critical dimension. By using more information about the overlaps between the sets $A(\mathbf{i}; \mathbf{j})$ for $\mathbf{j} \in \mathcal{I}_n$, we can say something slightly stronger and show the following:

Proposition 5.1. *Assume $M = N$ and $\mathbf{p} = (1/N, \dots, 1/N)$. Then $\mathcal{H}^s(F) = 0$ almost surely with respect to $\mu_{\mathbf{p}}$, where $s = -\log N / \log r$.*

(By this, we mean that there is a set of ω's of full measure under $\mu_{\mathbf{p}}$ for which $\mathcal{H}^s(F(\omega)) = 0$; note that it is not obvious *a priori* that the function $\omega \mapsto \mathcal{H}^s(F(\omega))$ is measurable. We therefore bound it by a function whose measurability is not in doubt and which takes the value 0 for $\mu_{\mathbf{p}}$-almost all ω.)

As a consequence of Proposition 5.1, we point out that when $M = N$ and $r = 1/N$, our construction generates a random subset of the interval which has Lebesgue measure zero but full Hausdorff dimension one almost surely.

Recall the definition of $a_{\mathbf{i}}$ from (4.1). We also continue to use Z_n for the number of basic intervals at level n in the construction of C needed to cover F. We first prove the following dichotomy.

Lemma 5.2. *The following statements hold.*

    (i)   *If $M \leq N$, then $E_{\mathbf{p}}(Z_n)/N^n \to 0$.*

    (ii)   *If $M > N$, then $E_{\mathbf{p}}(Z_n)/N^n \to \gamma$, for some number $\gamma \in (0, 1)$.*

Proof. By symmetry, $a_{\mathbf{i}}$ is the same for all $\mathbf{i} \in \Omega_n$, so we just consider the label sequence $\mathbf{i} = 1^n$. Set $\pi_n := a_{1^n}$. Then, by reasoning similar to that in Remark 4.1, π_n satisfies the recursion

$$\pi_n = \sum_{k=0}^{M} \binom{M}{k} \left(\frac{1}{N}\right)^k \left(1 - \frac{1}{N}\right)^{M-k} \left(1 - (1 - \pi_{n-1})^k\right)$$

$$= 1 - \left(1 - \frac{\pi_{n-1}}{N}\right)^M.$$

Clearly, π_n is decreasing in n, so its limit as $n \to \infty$ exists. Denote this limit by γ; then γ is a root of the equation

$$x = 1 - \left(1 - \frac{x}{N}\right)^M =: h(x).$$

Observe that

$$h'(x) = \frac{M}{N}\left(1 - \frac{x}{N}\right)^{M-1}, \qquad h''(x) = -\frac{M(M-1)}{N^2}\left(1 - \frac{x}{N}\right)^{M-2}.$$

This shows that h is strictly concave on $[0, 1]$, and if $M \leq N$, then $h' \leq 1$ on $[0, 1]$. Since furthermore, $h(0) = 0$, it follows that when $M \leq N$, h has a unique fixed point $x = 0$, and hence $\pi_n \to 0$ as $n \to \infty$. Part (i) of the lemma now follows since

$$E_{\mathbf{p}}(Z_n) = \sum_{\mathbf{i} \in \Omega_n} a_{\mathbf{i}} = N^n a_{1^n} = N^n \pi_n. \qquad (5.1)$$

On the other hand, if $M > N$, then $h'(0) = M/N > 1$ whereas $h(1) < 1$, so h has an additional fixed point $\gamma \in (0, 1)$. In this case, the fixed point $x = 0$ is repelling, while the fixed point γ is attracting. Thus, $\pi_n \to \gamma$ as $n \to \infty$, and part (ii) of the lemma follows by (5.1). ∎

We also need the following standard result from probability, of which we omit the proof.

Lemma 5.3. *Let $\{X_n\}_{n \in \mathbb{N}}$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. Suppose $E(X_n) \to 0$ as $n \to \infty$, and for all $n \in \mathbb{N}$ and all $\omega \in \Omega$ we have $0 \leq X_{n+1}(\omega) \leq X_n(\omega)$. Then $X_n \to 0$ almost surely.*

Proof of Proposition 5.1. Given $\delta > 0$, choose n_0 so large that $r^{n_0} < \delta$. Then, since $F(\omega)$ may be covered by $Z_n(\omega)$ intervals of length r^n, we have for all $n \geq n_0$ and all $\omega \in \Omega$

$$\mathcal{H}^s_\delta(F(\omega)) \leq r^{sn} Z_n(\omega) = N^{-n} Z_n(\omega).$$

(Recall $s = -\log N / \log r$). Note that $Z_{n+1}(\omega) \leq N Z_n(\omega)$ for all ω. Therefore, $N^{-n} Z_n$ is decreasing, and the limit in the following inequality exists for all ω:

$$\mathcal{H}^s_\delta(F(\omega)) \leq \lim_{n \to \infty} N^{-n} Z_n(\omega) =: Y(\omega).$$

As this holds for all $\delta > 0$, we have $\mathcal{H}^s(F(\omega)) \leq Y(\omega)$ for all $\omega \in \Omega$. Furthermore,

$$E_{\mathbf{p}}(N^{-n} Z_n) = \pi_n \to 0 \qquad \text{as } n \to \infty$$

by Lemma 5.2 (i). This implies $\mu_{\mathbf{p}}(Y = 0) = 1$ by applying Lemma 5.3 with $X_n = N^{-n} Z_n$, $P = \mu_{\mathbf{p}}$ and $(\Omega, \mathcal{F})$, completing the proof. ∎

Remark 5.4. By a very similar argument, we can show also that $\mathcal{H}^{s'}(F) = 0$ almost surely when $M < N$, where $s' = -\log M / \log r$. Thus, at least in the symmetric case of Corollary 2.4, we also have that the Hausdorff measure of F in the critical dimension is zero. Unfortunately, we do not know how to prove this, or whether it is in fact true, in the non-symmetric case.

While we did not use Lemma 5.2 (ii) in the above proof, it does serve to inform the following conjecture:

Conjecture 5.5. *When $M > N$ and $\mathbf{p} = (1/N, \dots, 1/N)$, $\mathcal{H}^s(F) > 0$ almost surely with respect to $\mu_{\mathbf{p}}$, where $s = -\log N / \log r$.*

Remark 5.6. When $N = 2$ and $M \geq 4$, something stronger than Conjecture 5.5 is true: With $\mu_{\mathbf{p}}$-probability 1, F contains a similar copy of the full Cantor set C (or contains an interval in case $r = 1/2$), and F is even equal to all of C with positive probability. As pointed out in [2], this is a consequence of [3, Corollary 6.2]. The argument there is easily extended to obtain the same conclusion for arbitrary $N \geq 2$ and $M \geq 2N$. This still leaves a gap in our knowledge: What happens when $N < M < 2N$?

Remark 5.7. In view of Proposition 5.1, it is natural to ask whether in the case $M = N$ there exists a gauge function φ (that is, a strictly increasing continuous function $\varphi : [0, \infty) \to [0, \infty)$ with $\varphi(0) = 0$) such that

$$0 < \mathcal{H}^\varphi(F(\omega)) < \infty \qquad \text{for } \mu_{\mathbf{p}}-\text{ almost every } \omega, \tag{5.2}$$

where for a set E,

$$\mathcal{H}^{\varphi}(E) := \lim_{\delta \searrow 0} \inf \left\{ \sum_i \varphi(|U_i|) : \{U_i\} \text{ is a } \delta\text{-cover of } E \right\}.$$

For instance, for $N = M = 2$, we can show that

$$\frac{1}{1+n} \le \pi_n \le \frac{4}{4+n} \qquad \forall\, n \ge 0.$$

We expect that similar inequalities hold in general when $N = M$. These inequalities, together with $E_{\mathbf{p}}(Z_n) = N^n \pi_n$ suggest that a good candidate is the function $\varphi(x) = x^s \log(1/x)$, where $s = -\log N / \log r$. Unfortunately, we have been unable to prove either inequality in (5.2) for this φ.

6. Deterministic case

In this section, we consider a simple deterministic scheme for labeling the tree, and compute the Hausdorff dimension of the resulting subsets of C. For the sake of simplicity, we take $N = M = 2$ and denote the maps of the IFS by f_0 and f_1. For any integer $m \ge 2$ let $g_m : \mathcal{I}^* \to \{0, 1\}$ be defined by

$$g_m(\mathbf{i}) = \begin{cases} 1 & \text{if } \kappa(\mathbf{i}) \equiv 0 \pmod{m}, \\ 0 & \text{otherwise.} \end{cases}$$

Then define the subset $F_m \subset C$ by

$$F_m := \left\{ x = \lim_{n \to \infty} f_{g_m(\mathbf{i}|_1)} \circ f_{g_m(\mathbf{i}|_2)} \cdots \circ f_{g_m(\mathbf{i}|_n)}([0, 1]) : \mathbf{i} \in \mathcal{I} \right\}.$$

Essentially, F_m is generated by the binary tree for which every m^{th} edge is labeled 1 while all others are labeled 0, see Figure 3. In the notation of the previous sections, $F_m = F(\omega)$ where $\omega = (0^{m-1}1)^{\infty}$.

Observe that $F_2 = C$, since g_2 is precisely the usual labeling of the full binary tree. By contrast, for $m \ge 3$, F_m has an interesting connection to golden mean-like shifts of finite type.

For $L \in \mathbb{N}$, let $X_L \subseteq \{0, 1\}^{\mathbb{N}}$ be the one-sided subshift of finite type with forbidden words list

$$\mathcal{F}_L := \{10^k 1 : 0 \le k \le L - 1\}.$$

Thus, a sequence $\mathbf{i} \in \{0, 1\}^{\mathbb{N}}$ lies in X_L if and only if any two consecutive 1's in $\mathbf{i}$ are separated by a string of at least L consecutive 0's. It is well known that X_L has entropy $\log \rho_L$, where $\rho_L \in (1, 2)$ is the positive real root of $1 + x^L = x^{L+1}$ (see [15, Exercise 4.3.7]). In particular, $\rho_1 = (1 + \sqrt{5})/2$ is the golden ratio.
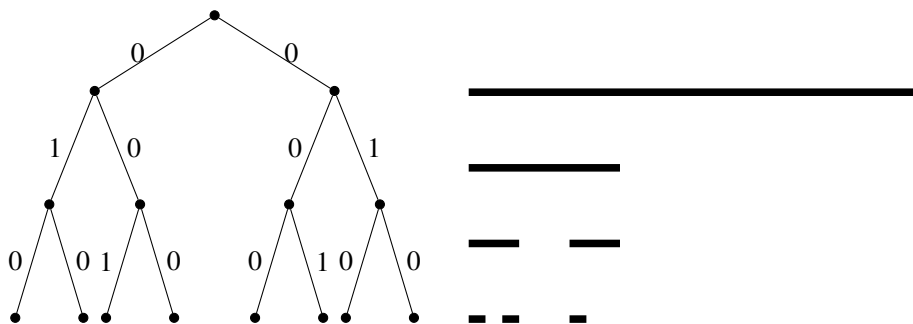
Figure 3. An example tree with $m = 3$ (left) and the third level approximation of F_3 (right).

Proposition 6.1. *For an integer $m \geq 3$, let L be the integer such that*

$$m \in \{2^{L+1} - 1, 2^{L+1}, \ldots, 2^{L+2} - 2\}.$$

Then

$$\pi^{-1}[F_m] = \{\mathbf{j} = (j_1, j_2, \ldots) \in X_L : j_1 = j_2 = \cdots = j_L = 0\}, \qquad (6.1)$$

where π is the natural projection of the symbolic space onto the Cantor set C, i.e.,

$$\pi(\mathbf{j}) := \lim_{n \to \infty} f_{j_1} \circ \cdots \circ f_{j_n}(0) \text{ for } \mathbf{j} = (j_1, j_2, \ldots) \in \{0, 1\}^{\mathbb{N}}.$$

Consequently,

$$\dim_H F_m = \frac{\log \rho_L}{-\log r}. \qquad (6.2)$$

(Recall that r is the common contraction ratio of the maps f_0 and f_1.)

Proof. First, construct a directed graph G whose vertices are labeled $0, 1, \ldots, m - 1$ to represent the residual classes mod m. Then draw an edge from each vertex j to $2j + 1$ mod m and $2j + 2$ mod m (see Figure 4 for the case $m = 6$). Notice in particular that

$$2(m - 1) + 1 \equiv m - 1 \pmod{m} \quad \text{and} \quad 2(m - 1) + 2 \equiv 0 \pmod{m}.$$

Thus, vertex $m - 1$ admits a self loop, as well as an edge connected to vertex 0. Furthermore, for any $k \leq L + 1$ and $2^k - 1 \leq j \leq \min\{2^{k+1} - 2, m - 1\}$, there is a walk of length k from vertex 0 to vertex j. This becomes clear when drawing G as a tree rooted at 0 and omitting incoming edges, as shown in Figure 5.

Now, by construction, every word in $\pi^{-1}[F_m]$ can be represented as an infinite walk in G beginning at either vertex 1 or 2. That is, any step passing through vertices $1, \ldots, m - 1$ records a 0, while a step through vertex 0 records a 1.
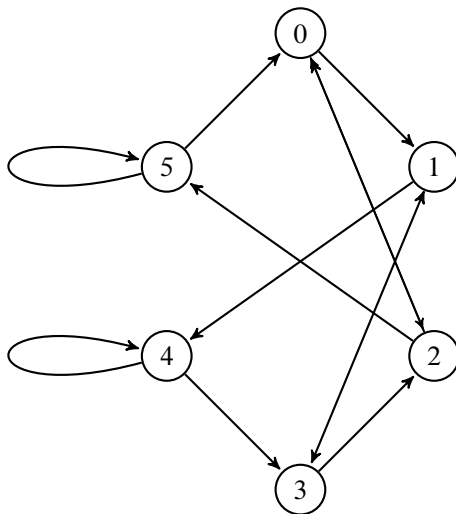
Figure 4. The directed graph G, shown for $m = 6$. Each vertex j has outgoing edges to $2j + 1$ mod m and $2j + 2$ mod m. Vertex $m - 1$ always has a self-loop and an outgoing edge to 0.

Since $2^{L+1} - 1 \leq m \leq 2^{L+2} - 2$, we have a minimum of $L + 1$ steps to walk from vertex 0 back to itself in G. Thus, no word in $\mathcal{F}_L$ can appear as a sub-word of any word in $\pi^{-1}[F_m]$. Consequently, $\pi^{-1}[F_m] \subseteq X_L$.

To show the other inclusion, we need to show that for every integer $l \geq L + 1$, there exists a walk in G from vertex 0 back to itself of length l (without passing through 0 along the way). This is clear for $l = L + 1$, so assume $l > L + 1$. If $m = 2^{L+1} - 1$, we take the L-step walk

$$0 \to 2 \to 6 \to \cdots \to 2^{L+1} - 2 = m - 1,$$

then repeat the self loop at vertex $m - 1$ for $l - L - 1$ times, and finally we take $m - 1 \to 0$ for the last step. On the other hand, if $m \geq 2^{L+1}$, then as previously discussed, we may reach vertex $m - 1$ in $L + 1$ steps, repeat the self loop for $l - L - 2$ times, then take $m - 1 \to 0$ as the last step. Therefore, every sequence in X_L beginning with 0^L lies in $\pi^{-1}[F_m]$.

This gives (6.1). Finally, since the IFS $\{f_0, f_1\}$ satisfies the open set condition, it follows from standard arguments that the Hausdorff dimension of F_m is given by the entropy of X_L over $\log r^{-1}$ (see [19, Lemma 16.4.2] for example), yielding (6.2). ∎

It may be of interest to compare the value $\dim_H F_m = -\log \rho_L / \log r$ with the bounds from Theorem 2.2 and Corollary 2.5 for the random case with $N = M = 2$ and $p = 1/m$.
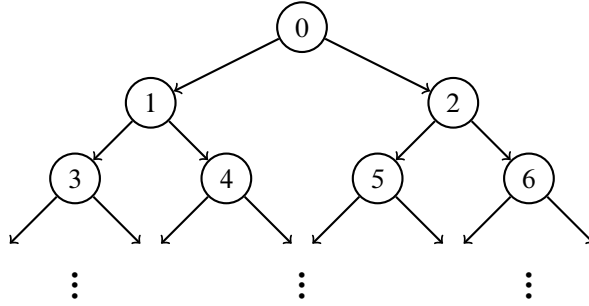
Figure 5. Each vertex j for $2^k - 1 \le j \le 2^{k+1} - 2$ may be reached in exactly k steps.

m	p	lower bound	$\dim_H F_m$	upper bound
2	1/2	.631	.631	.631
3	1/3	.535	.438*	.618
4	1/4	.428	.438	.599
6	1/6	.296	.438	.569
7	1/7	.256	.348	.557
14	1/14	.130	.348	.503
15	1/15	.121	.293	.498
30	1/30	.061	.256	.450

Table 1. Comparison of $\dim_H F_m$ with the upper and lower bounds from Theorem 2.2 and Corollary 2.5 with $p = 1/m$ and $r = 1/3$, for select values of m. Note that $\dim_H F_m$ is constant on each interval $2^{L+1} - 1 \le m \le 2^{L+2} - 2$, for $L \ge 1$. It appears that only the value for $m = 3$ is outside the bounds.

We have made such a comparison in Table 1. Although there is no reason why the dimension of F_m should lie between the bounds of Theorem 2.2 (since F_m is but one of uncountably many realizations of the random subset $F(\omega)$), this does appear to be the case for all $m \ge 4$.

References

[1] M. F. Barnsley, J. E. Hutchinson, and Ö. Stenflo, *V*-variable fractals: Fractals with partial self similarity. *Adv. Math.* **218** (2008), no. 6, 2051–2088 Zbl 1169.28006 MR 2431670

[2] I. Benjamini, O. Gurel-Gurevich, and B. Solomyak, Branching random walk with exponentially decreasing steps, and stochastically self-similar measures. *Trans. Amer. Math. Soc.* **361** (2009), no. 3, 1625–1643 Zbl 1162.60023 MR 2457411

[3] I. Benjamini and H. Kesten, Percolation of arbitrary words in $\{0, 1\}^{\mathbb{N}}$. *Ann. Probab.* **23** (1995), no. 3, 1024–1060 Zbl 0832.60095 MR 1349161

[4] C. Chen, A class of random Cantor sets. *Real Anal. Exchange* **42** (2017), no. 1, 79–120 Zbl 1384.28012 MR 3702557

[5] M. Dekking and K. Simon, On the size of the algebraic difference of two random Cantor sets. *Random Structures Algorithms* **32** (2008), no. 2, 205–222 Zbl 1141.60004 MR 2387557

[6] K. Falconer, Random fractals. *Math. Proc. Cambridge Philos. Soc.* **100** (1986), no. 3, 559–582 Zbl 0623.60020 MR 0857731

[7] K. Falconer, Dimensions and measures of quasi self-similar sets. *Proc. Amer. Math. Soc.* **106** (1989), no. 2, 543–554 Zbl 0683.58034 MR 0969315

[8] K. Falconer, *Techniques in fractal geometry*. John Wiley & Sons, Ltd., Chichester, 1997 Zbl 0869.28003 MR 1449135

[9] K. Falconer, *Fractal geometry: Mathematical foundations and applications*, Second edition. John Wiley & Sons, Inc., Hoboken, NJ, 2003 Zbl 1060.28005 MR 2118797

[10] S. Graf, Statistically self-similar fractals. *Probab. Theory Related Fields* **74** (1987), no. 3, 357–392 Zbl 0591.60005 MR 0873885

[11] J. E. Hutchinson, Fractals and self-similarity. *Indiana Univ. Math. J.* **30** (1981), no. 5, 713–174 Zbl 0598.28011 MR 0625600

[12] E. Järvenpää, M. Järvenpää, A. Käenmäki, H. Koivusalo, Ö. Stenflo, and V. Suomala, Dimensions of random affine code tree fractals. *Ergodic Theory Dynam. Systems* **34** (2014), 854–875 Zbl 1298.28017 MR 3199797

[13] H. Koivusalo, Dimension of uniformly random self-similar fractals. *Real Anal. Exchange* **39** (2014), no. 1, 73–90 Zbl 1303.28013 MR 3261900

[14] P. Larsson, L'ensemble différence de deux ensembles de Cantor aléatoires. *C. R. Acad. Sci. Paris Sér. I Math.* **310** (1990), no. 10, 735–738 Zbl 0707.60074 MR 1055239

[15] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995 Zbl 1106.37301 MR 1369092

[16] Y.-Y. Liu and J. Wu, A dimensional result for random self-similar sets. *Proc. Amer. Math. Soc.* **130** (2002), no. 7, 2125–2131 Zbl 0989.60013 MR 1896049

[17] R. D. Mauldin and S. C. Williams, Random recursive constructions: asymptotic geometric and topological properties. *Trans. Amer. Math. Soc.* **295** (1986), no. 1, 325–346 Zbl 0625.54047 MR 0831202

[18] P. Shmerkin and V. Suomala, Sets which are not tube null and intersection properties of random measures. *J. Lond. Math. Soc. (2)* **91** (2015), no. 2, 405–422 Zbl 1317.28004 MR 3355108

[19] M. Urbański, M. Roy and S. Munday, *Non-invertible dynamical systems*, Vol. 2. De Gruyter Expositions in Mathematics 69, De Gruyter, Berlin, 2022 Zbl 1505.37003

Pieter Allaart (corresponding author)
Department of Mathematics, University of North Texas, 1155 Union Cir #311430, Denton, TX 76203-5017, USA; allaart@unt.edu

Taylor Jones
Applied Mathematics and Statistics Department, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218-2683, USA; rjone197@jh.edu