HAUSDORFF DIMENSION FOR FRACTALS INVARIANT UNDER THE MULTIPLICATIVE INTEGERS

RICHARD KENYON, YUVAL PERES, AND BORIS SOLOMYAK

Abstract. We consider subsets of the (symbolic) sequence space that are invariant under the action of the semigroup of multiplicative integers. A representative example is the collection of all 0-1 sequences (x_k) such that $x_k x_{2k} = 0$ for all k. We compute the Hausdorff and Minkowski dimensions of these sets and show that they are typically different. The proof proceeds via a variational principle for multiplicative subshifts.

1. Introduction

Central objects in symbolic dynamics and the theory of fractals are shifts of finite type, and more generally, closed subsets of the symbolic space $\Sigma_m := \{0, \ldots, m-1\}^{\mathbb{N}}$ that are invariant under the shift $\sigma(x_1, x_2, x_3, \ldots) = (x_2, x_3, \ldots)$. (we refer to them as "subshifts" for short). To a subset Ω of Σ_m we can associate a subset of $[0,1]$ by considering the collection of all reals whose base m digit sequences belong to Ω. Subshifts then correspond to closed subsets of $[0,1]$ invariant under the map $x \mapsto mx \pmod 1$. It is known [7] that all such sets have the Hausdorff dimension equal to the Minkowski (box-counting) dimension, which is equal to $(\log m)^{-1}$ times the topological entropy of σ on Ω.

Note that shift-invariance implies invariance under the action of the **semigroup of additive positive integers.** In contrast, in this paper we consider subsets of Σ_m and the corresponding fractals in $[0,1]$, which arise from the action of the **semigroup of multiplicative integers.** Namely, given a subset $\Omega \subset \Sigma_m$ and an integer $q \geq 2$, we let

$$X_\Omega = X_\Omega^{(q)} := \left\{ \omega = (x_k)_{k=1}^\infty \in \Sigma_m : \ (x_{iq^\ell})_{\ell=0}^\infty \in \Omega \ \text{ for all } i, \ q \nmid i \right\} \tag{1}$$

Date: July 20, 2011.

The research of R. K. was supported in part by NSF. The research of B. S. was partially supported by the NSF grant DMS-0968879.

and consider the corresponding subset of $[0,1]$:

$$\Xi_\Omega := \Big\{x = \sum_{k=1}^{\infty} x_k m^{-k} :\ (x_k)_1^\infty \in X_\Omega\Big\}. \tag{2}$$

If Ω is shift-invariant, then X_Ω is invariant under the action of multiplicative integers:

$$(x_k)_1^\infty \in X_\Omega \ \Rightarrow\ (x_{rk})_{k=1}^\infty \in X_\Omega \ \text{ for all } r \in \mathbb{N}.$$

If Ω is a shift of finite type, we refer to X_Ω (and Ξ_Ω) as the "multiplicative shift of finite type."

Our interest in these sets was prompted by work of Ai-Hua Fan, Lingmin Liao and Jihua Ma [5] who computed the Minkowski dimension of the "multiplicative golden mean shift"

$$\Xi_g := \Big\{x = \sum_{k=1}^{\infty} x_k 2^{-k} :\ x_k \in \{0,1\},\ x_k x_{2k} = 0 \ \text{ for all } k\Big\} \tag{3}$$

and raised the question of computing its Hausdorff dimension. They showed that the Minkowski dimension is

$$\dim_M(\Xi_g) = \sum_{k=1}^{\infty} \frac{\log_2 F_{k+1}}{2^{k+1}} = 0.82429\ldots, \tag{4}$$

where F_k is the k-th Fibonacci number: $F_1 = 1,\ F_2 = 2, F_{k+1} = F_{k-1} + F_k$. As a special case of our results we obtain the Hausdorff dimension $\dim_H(\Xi_g)$.

Proposition 1.1. *We have*

$$\dim_H(\Xi_g) = -\log_2 p = 0.81137\ldots, \quad \textit{where } p^3 = (1-p)^2, \quad 0 < p < 1 \tag{5}$$

Thus, $\dim_H(\Xi_g) < \dim_M(\Xi_g)$.

Proposition 1.1 will follow from a more general result, Theorem 1.3 below. For an exposition which focuses on the set Ξ_g see [8].

In order to visualize the set Ξ_g we show the set $\widetilde{\Xi}_g$ in Figure 1, which is obtained from Ξ_g by the transformation

$$\sum_{k=1}^{\infty} x_k 2^{-k} \mapsto \Big(\sum_{k=1}^{\infty} x_{2k-1} 2^{-k}, \sum_{k=1}^{\infty} x_{2k} 2^{-k}\Big).$$

It is easy to see that this transformation doubles the Minkowski and Hausdorff dimensions.

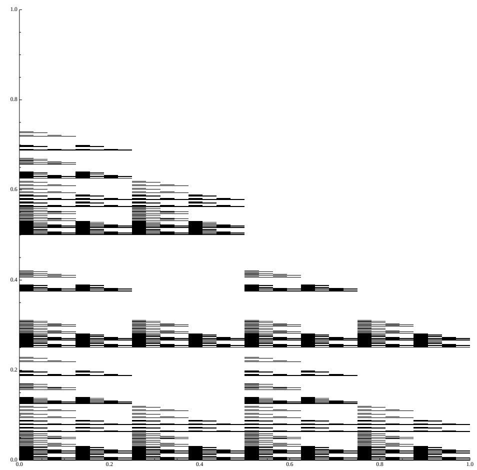

FIGURE 1. Approximation of the set $\widetilde{\Xi}_g$.

The figure resembles pictures of self-affine carpets, see [1, 10], for which the Hausdorff dimension is often less than the Minkowski dimension. In fact, our proof bears some similarities with those of [1, 10] as well. An example of a self-affine set is shown in Figure 2.

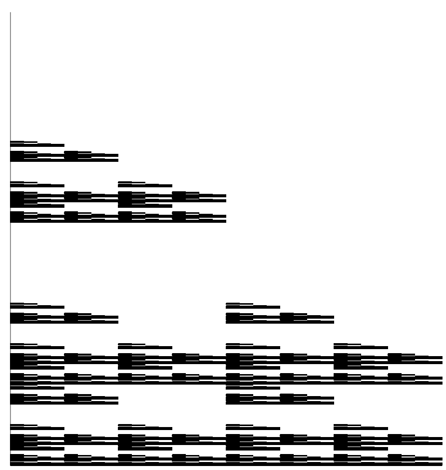

FIGURE 2. Approximation of a self-affine set. Letting $M(x, y) = (\frac{x}{2}, \frac{y}{4})$, this set is $S = \left\{ \sum_{i \geq 0} M^i d_i : \, d_i \in \{(0,0),(1,0),(0,1),(1,1),(0,3)\} \right\}$.

The set Ξ_g is a representative example of a large family of sets for which we compute the dimension. Let $m \geq 2$ and let $A = (A(i,j))_{i,j=0}^{m-1}$ be a primitive (a non-negative matrix is primitive if some power is strictly positive) matrix with 0-1 entries. The usual (additive) shift of finite type determined by A is defined as

$$\Sigma_A := \{(x_k)_{k=1}^{\infty}:\ x_k \in \{0,\dots,m-1\},\ A(x_k, x_{k+1}) = 1,\ k \ge 1\}.$$

Instead, we fix an integer $q \ge 2$ and consider the multiplicative shift of finite type

$$X_A = X_A^{(q)} := \{(x_k)_{k=1}^{\infty}:\ x_k \in \{0,\dots,m-1\},\ A(x_k, x_{qk}) = 1,\ k \ge 1\}, \tag{6}$$

as well as the corresponding subset of the unit interval:

$$\Xi_A := \Big\{x = \sum_{k=1}^{\infty} x_k m^{-k},\ (x_k)_{k=1}^{\infty} \in X_A\}.$$

As is well-known, the dimensions of Ξ_A and X_A coincide, if we use the standard metric on the sequence space Σ_m:

$$\varrho\big((x_k),(y_k)\big) = m^{-\min\{n:\ x_n \ne y_n\}+1}$$

on the sequence space Σ_m; this is equivalent to restricting the covers of Ξ_A to those by m-adic intervals. Thus, in the rest of the paper we focus on the sets X_A.

In order to state our dimension result, we need the following elementary lemma.

Lemma 1.2. *Let $A = (A(i,j))_{i,j=0}^{m-1}$ be a primitive matrix, and $q > 1$. Then there exists a unique vector $(t_i)_{i=0}^{m-1}$ satisfying*

$$t_i^q = \sum_{j=0}^{m-1} A(i,j) t_j, \quad t_i > 1,\ i = 0,\dots,m-1. \tag{7}$$

Below we use logarithms to base m, denoted $\log_m$, and write $\overline{1}$ for the vector $(1,\dots,1)^T \in \mathbb{R}^m$.

Theorem 1.3. (i) *Let A be a primitive 0-1 matrix. Then the set X_A given by (6) satisfies*

$$\dim_H(X_A) = \frac{q-1}{q} \log_m \sum_{i=0}^{m-1} t_i, \tag{8}$$

where $(t_i)_{i=0}^{m-1}$ is from Lemma 1.2.

(ii) *The Minkowski dimension of X_A exists and equals*

$$\dim_M(X_A) = (q-1)^2 \sum_{k=1}^{\infty} \frac{\log_m(A^{k-1}\overline{1}, \overline{1})}{q^{k+1}}. \tag{9}$$

We have $\dim_H(X_A) = \dim_M(X_A)$ *if and only if* A *has* $\overline{1}$ *as an eigenvector (i.e. row sums of* A *are all equal).*

The formula for the Minkowski dimension is not difficult to prove; it is included for comparison.

1.1. **Variational principle for multiplicative subshifts.** We obtain Theorem 1.3 as a special case of a more general result. Let Ω be an arbitrary closed subset of Σ_m (it does not have to be shift-invariant), and define the sets X_Ω and Ξ_Ω by (1) and (2). We refer to X_Ω as a "multiplicative subshift." Precise statements are given in the next section; here we just describe the results.

We can view our set X_Ω as an infinite union of copies of Ω, starting at all positive integers i not divisible by q (denoted $q \nmid i$) and "sitting" along geometric progressions of ratio q. More precisely, denote $J_i = \{q^r i\}_{r=0}^{\infty}$ for $q \nmid i$ and let $x|J_i = (x_{q^r i})_{r=0}^{\infty}$. By definition (1),

$$x \in X_\Omega \iff x|J_i \in \Omega \text{ for all } i,\ q \nmid i. \tag{10}$$

In order to compute (or estimate) the Hausdorff dimension of a set, one usually has to equip it with a "good" measure and calculate the appropriate "Hölder exponent". For subshifts, "good" measures are ergodic invariant measures. For multiplicative subshifts, their role is played by measures obtained in the following construction, essentially as an infinite product of copies of a measure on Ω. Given a probability measure μ on Ω we set

$$\mathbb{P}_\mu[u] := \prod_{i \le |u|,\, q \nmid i} \mu[u|J_i], \tag{11}$$

where $[u]$ denotes the cylinder set of all sequences starting with u and

$$u|J_i = u_i u_{qi} \dots u_{q^r i}, \quad q^r i \le |u| < q^{r+1} i.$$

It is easy to verify that $\mathbb{P}_\mu$ is a Borel probability measure supported on X_Ω (see the next section for details).

For a probability measure $\mathbb{P}$, its Hausdorff dimension is defined by

$$\dim_H(\mathbb{P}) = \inf\{\dim_H(F) :\ F \text{ Borel},\ \mathbb{P}(F) = 1\},$$

and the pointwise dimension at x is given by

$$\dim_{\text{loc}}(\mathbb{P}, x) = \lim_{r \to 0} \frac{\log \mathbb{P}(B_r(x))}{\log r},$$

whenever the limit exists, where $B_r(x)$ denotes the open ball of radius r centered at x. We consider measures on the sequence space Σ_m; then

$$\dim_{\rm loc}(\mathbb{P}, x) = \lim_{n\to\infty} \frac{-\log \mathbb{P}\,[x_1^n]}{\log n}, \tag{12}$$

where $x_1^n = x_1 \dots x_n$ denotes the initial segment (prefix) of the sequence x. We prove that for any measure $\mathbb{P}_\mu$ defined above, the pointwise dimension exists and is constant $\mathbb{P}_\mu$-a.e., which is then equal to $\dim_H(\mathbb{P}_\mu)$ (see Proposition 2.3). This can be viewed as a multiplicative analog of the Shannon-McMillan-Breiman Theorem and the entropy formula for the dimension of an ergodic shift-invariant measure ν, namely, $\dim_H(\nu) = h(\nu)/\log m$ (see [2]). Further, we obtain the "Variational Principle for multiplicative subshifts," see Proposition 2.4. We can summarize this discussion with the following dictionary between the classical and multiplicative subshifts:

classical	**multiplicative**
subshift $\Upsilon \subset \Sigma_m$	set X_Ω
invariant ergodic measure ν on Υ	measure $\mathbb{P}_\mu$
Shannon-McMillan-Breiman Theorem	pointwise dimension of $\mathbb{P}_\mu$
$\dim_H(\nu) = h(\nu)/\log m$	dimension of $\mathbb{P}_\mu$
Variational Principle: $\dim_H(\Upsilon) = \sup\{\dim_H(\nu) :\ \nu \text{ is ergodic on } \Upsilon\}$	$\dim_H(X_\Omega) = \sup\{\dim_H(\mathbb{P}_\mu) :$ μ is a probability on $\Omega\}$

2. General result. Variational problem.

Let Ω be an arbitrary closed subset of Σ_m, and define the sets X_Ω and Ξ_Ω by (1) and (2). Our general theorem computes the Hausdorff and Minkowski dimensions of X_Ω (as discussed earlier, the dimensions of Ξ_Ω are the same as those of X_Ω).

Consider the **tree of prefixes** of the set Ω. It is a directed graph $\Gamma = \Gamma(\Omega)$ whose set of vertices is

$$V(\Gamma) = \mathrm{Pref}(\Omega) = \bigcup_{k=0}^{\infty} \mathrm{Pref}_k(\Omega),$$

where $\mathrm{Pref}_0(\Omega)$ has only one element, the empty word $\varnothing$, and

$$\mathrm{Pref}_k(\Omega) = \{u \in \{0, \dots, m-1\}^k,\ \Omega \cap [u] \neq \emptyset\}.$$

There is a directed edge from a prefix u to a prefix v if $v = ui$ for some $i \in \{0, \dots, m-1\}$. In addition, there is an edge from $\varnothing$ to every $i \in \mathrm{Pref}_1(\Omega)$. Clearly, $\Gamma(\Omega)$ is a tree, and it has the outdegree bounded by m. Note that if Ω is shift-invariant, then the set $\mathrm{Pref}(\Omega)$ coincides with the set of allowed (admissible) words in Ω (sometimes referred to as the **language** of Ω).

The next lemma generalizes Lemma 1.2.

Lemma 2.1. *Let $\Gamma = (V, E)$ be a directed graph (finite or infinite) with the outdegree bounded by $M < \infty$, such that from each vertex there is at least one outgoing edge. Let $q > 1$. Then there exists a unique vector $\bar{t} \in [1, M^{\frac{1}{q-1}}]^V$ such that*

$$t_v^q = \sum_{vw \in E} t_w, \quad v \in V. \tag{13}$$

It is clear that Lemma 1.2 is a special case, with Γ being the directed graph with the incidence matrix A.

Note that we only claim uniqueness of solutions in the given range. In fact, uniqueness of positive solutions holds if we assume a priori bounds from zero and infinity; without this assumption there may be infinitely many solutions on an infinite graph.

Theorem 2.2. *Let $\Omega \subset \Sigma$, and let $\bar{t}$ be the vector from Lemma 2.1 corresponding to the tree of prefixes $\Gamma(\Omega)$. Then*

(i)

$$\dim_H(X_\Omega) = (q-1)\log_m t_\emptyset; \tag{14}$$

(ii)

$$\dim_M(X_\Omega) = (q-1)^2 \sum_{k=1}^{\infty} \frac{\log_m |\mathrm{Pref}_k(\Omega)|}{q^{k+1}}. \tag{15}$$

We have $\dim_H(X_\Omega) = \dim_M(X_\Omega)$ if and only if the tree of prefixes is spherically symmetric*, i.e. for every $k \in \mathbb{N}$, all prefixes of length k have the same (equal) number of continuations in $\mathrm{Pref}_{k+1}(\Omega)$.*

Observe that Theorem 1.3 is a special case of Theorem 2.2: For part (i), we note that for a shift of finite type Σ_A the graph $\Gamma(\Sigma_A)$ has the property that the tree of descendants of a prefix $u = u_1 \dots u_k$ depends only on the last symbol u_k. Denote by T_i this tree, which has $u_k = i$ as its root vertex, for $i = 0, \dots, m-1$, and let t_i be the solution of the system of equations (13) evaluated at the root. Here we use Lemma 2.1, with the uniqueness statement. Then we obtain from (13) that the vector $(t_i)_{i=0}^{m-1}$ satisfies (7). Finally, note that $t_\varnothing^q = \sum_{i=0}^{m-1} t_i$ by (13), hence (14) reduces to (8).

For part (ii), we just note that $(A^{k-1}\overline{1}, \overline{1})$ is the number of allowed words of length k in the shift of finite type Σ_A.

2.1. **Scheme of the proof. Statement of the Variational Principle.** Recall (11) that, given a probability measure μ on Ω we define a measure on X_Ω by

$$\mathbb{P}_\mu[u] := \prod_{i \le n,\, q \nmid i} \mu[u|J_i], \quad \text{where } |u| = n \text{ and } J_i = \{q^r i\}_{r=1}^\infty. \tag{16}$$

This is a well-defined pre-measure on the semi-algebra of cylinder sets. Indeed, we have $\mathbb{P}_\mu[i] = \mu[i]$ for $i = 0, \dots, m-1$, and for $n+1 = q^r i,\ q \nmid i$,

$$\frac{\mathbb{P}_\mu[u_1 \dots u_n u_{n+1}]}{\mathbb{P}_\mu[u_1 \dots u_n]} = \frac{\mu[u_i u_{qi} \dots u_{q^r i}]}{\mu[u_i u_{qi} \dots u_{q^{r-1} i}]},$$

whence

$$\mathbb{P}_\mu[u_1 \dots u_n] = \sum_{j=0}^{m-1} \mathbb{P}_\mu[u_1 \dots u_n j].$$

The extension of $\mathbb{P}_\mu$ is a Borel measure supported on X_Ω, since Ω is a closed subset of Σ_m and hence

$$\Omega = \bigcap_{k=1}^\infty \bigcup_{u \in \mathrm{Pref}_k(\Omega)} [u].$$

Observe that (16) is not the only way to put a measure on X_Ω: we could make the measure $\mu = \mu_i$ in (16) depend on i; however, this is not necessary for the purpose of computing the Hausdorff dimension.

We compute the Hausdorff dimension $\dim_H(\mathbb{P}_\mu)$, which yields a lower bound on $\dim_H(X_\Omega)$. In order to state the result, we need to introduce some notation.

For $k \ge 1$ let α_k be the partition of Ω into cylinders of length k:

$$\alpha_k = \{\Omega \cap [u] :\ u \in \mathrm{Pref}_k(\Omega)\} = \{\Omega \cap [u] :\ u \in \{0, \dots, m-1\}^k,\ \Omega \cap [u] \ne \emptyset\}.$$

For a measure μ on Σ_m and a finite partition α, denote by $H_m^\mu(\alpha)$ the μ-entropy of the partition, with base m logarithms:

$$H_m^\mu(\alpha) = -\sum_{C \in \alpha} \mu(C) \log_m \mu(C).$$

Now define

$$s(\Omega,\mu) := (q-1)^2 \sum_{k=1}^{\infty} \frac{H_m^{\mu}(\alpha_k)}{q^{k+1}}\,. \tag{17}$$

Proposition 2.3. *Let Ω be a closed subset of Σ_m and μ a probability measure on Ω. Then*

$$\dim_{\mathrm{loc}}(\mathbb{P}_\mu, x) = s(\Omega,\mu) \quad \text{for } \mathbb{P}_\mu\text{-a.e. } x \in X_\Omega. \tag{18}$$

Therefore, $\dim_H(\mathbb{P}_\mu) = s(\Omega,\mu)$, *and* $\dim_H(X_\Omega) \ge s(\Omega,\mu)$.

We also have the Variational Principle:

Proposition 2.4. *Let Ω be a closed subset of Σ_m. Then*

$$\dim_H(X_\Omega) = \sup_\mu \dim_H(\mathbb{P}_\mu) = \sup_\mu s(\Omega,\mu), \tag{19}$$

where the supremum is over Borel probability measures on Ω.

It is clear from (17) that the function $\mu \mapsto s(\Omega,\mu)$ is continuous on the compact space of probability measures with the w^*-topology. Thus, the supremum in (19) is actually a maximum. Let

$$s(\Omega) := \max\{s(\Omega,\mu) :\ \mu \text{ is a probability on } \Omega\}. \tag{20}$$

We call a measure μ for which $s(\Omega) = s(\Omega,\mu)$ an **optimal measure**. The next theorem characterizes such measures.

Proposition 2.5. *Let Ω be a closed subset of Σ_m and let $\bar{t}$ be the solution of the system of equations (13) for the tree of prefixes of Ω. For any $k \ge 1$ and $u \in \mathrm{Pref}_k(\Omega)$ let*

$$\mu[u] := \prod_{j=1}^{k} \frac{t_{u_1\ldots u_j}}{t^q_{u_1\ldots u_{j-1}}}\,. \tag{21}$$

This defines a probability measure μ on Ω. Moreover,

(i) *μ is the unique optimal measure;*

(ii) $s(\Omega,\mu) = (q-1)\log_m t_{\varnothing}$.

Combining Propositions 2.4 and 2.5 yields part (i) of Theorem 2.2.

In the case when Ω is a shift of finite type, the optimal measure turns out to be Markov.

Corollary 2.6. *Let A be a primitive $m\times m$ 0-1 matrix and Σ_A the corresponding shift of finite type. Let $\bar{t} = (t_i)_{i=0}^{m-1}$ be the solution of the system of equations (7). Then the unique optimal measure on Σ_A is Markov, with the vector of initial probabilities $\mathbf{p} = (\sum_{i=0}^{m-1} t_i)^{-1}\bar{t}$ and the matrix of transition probabilities*

$$(p_{ij})_{i,j=0}^{m-1} \text{ where } p_{ij} = \frac{t_j}{t_i^q} \quad \text{if } A(i,j) = 1.$$

3. Examples

Example 3.1 (golden mean)**.** *Let $q = 2, m = 2$, and $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. Then $\Xi_A = \Xi_g$, the multiplicative golden mean shift from (3).*

The system of equations (7) reduces to

$$t_0^2 = t_0 + t_1,\ t_1^2 = t_0,$$

which immediately implies $t_1^3 = t_1 + 1$. According to Corollary 2.6, the optimal measure μ on Σ_A is Markov, with initial probability of 0 equal to $p = t_0/(t_0+t_1) = t_0^{-1} = t_1^{-2}$, and the initial probability of 1 equal to $1-p = t_1/(t_0+t_1) = t_1^{-3}$, whence $p^3 = (1-p)^2$. The matrix of transition probabilities is $\begin{bmatrix} p & 1-p \\ 1 & 0 \end{bmatrix}$. Then, by (8),

$$\dim_H(\Xi_g) = (1/2)\log_2(t_0+t_1) = -\log_2 p,$$

which proves Proposition 1.1. □

Example 3.2 (Tribonacci)**.** *Let $q = 2,\ m = 3$, and $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Then*

$$X_A = \{(x_k)_1^\infty \in \{0,1,2\}^{\mathbb{N}} :\ x_k = 1 \ \Rightarrow\ x_{2k} = 0,\ x_k = 2 \ \Rightarrow\ x_{2k} = 1\}.$$

We have

$$\dim_H(X_A) = 4\log_3 t \approx 0.726227, \quad \text{where } t^4 - t - 1 = 0, \tag{22}$$

and

$$\dim_M(X_A) = \sum_{k=1}^{\infty} \frac{\log_3 T_{k-1}}{2^{k+1}} \approx 0.75373,$$

where $T_0 = 3,\ T_1 = 5,\ T_2 = 9,\ T_{k+2} = T_{k-1} + T_k + T_{k+1}$.

To verify (22), we note that the equations (7) in this case are

$$t_0^2 = t_0 + t_1 + t_2,\ t_1^2 = t_0,\ t_2^2 = t_1,$$

whence $t_2^8 = t_2^4 + t_2^2 + t_2$. Thus, $t = t_2$ satisfies $t^7 = t^3 + t + 1$, and since $t^7 - t^3 - t - 1 = (t^4 - t - 1)(t^3 + 1)$, Theorem 1.3(i) yields the formula for the Hausdorff dimension. The optimal measure is Markov, with the matrix of transition probabilities equal to $\begin{bmatrix} t^{-4} & t^{-6} & t^{-7} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, and the vector of initial probabilities (t^{-4}, t^{-6}, t^{-7}).

Example 3.3 (2-step Markov)**.** *Let* $q = 2$, $m = 2$, *and*

$$X := \left\{(x_k)_1^\infty \in \{0,1\}^{\mathbb{N}} :\ x_k x_{2k} x_{4k} = 0,\ k \ge 1\right\}.$$

Then $X = X_\Omega$ *where* Ω *is the shift of finite type on the alphabet* $\{0, 1\}$ *with the only forbidden 3-letter word* 111*.*

The graph $\Gamma(\Omega)$ has the property that the tree of descendants of a prefix $u = u_1 \dots u_k$ of length $|u| \ge 2$ depends only on the last pair of symbols $u_{k-1}u_k = ij$. Denote by T_{ij} this tree, for $i, j \in \{0, 1\}$, and let t_{ij} be the solution of the system of equations (13) evaluated at its root (we are using Lemma 2.1 with uniqueness here). Then (13) on $\Gamma(\Omega)$ yields

$$\begin{aligned} t_{00}^2 &= t_{00} + t_{01} \\ t_{01}^2 &= t_{10} + t_{11} \\ t_{10}^2 &= t_{00} + t_{01} \\ t_{11}^2 &= t_{10}, \end{aligned}$$

and also $t_\varnothing = t_{00}$. Denoting $z = t_{11}$ we obtain, after a simple computation, that $(z^4 - z^2)^2 = z^2 + z$ whence

$$z^7 - 2z^5 + z^3 - z - 1 = 0.$$

Note also that $t_{00} = t_{11}^2 = z^2$. Thus, by (14),

$$\dim_H(X) = 2 \log_2 z \approx 0.956651.$$

The Minkowski dimension in this example is

$$\dim_M(X) = \sum_{k=1}^{\infty} \frac{\log_2 R_{k-1}}{2^{k+1}} \approx 0.961789,$$

where R_j is the number of allowed sequences of length j: $R_1 = 2, R_2 = 4, R_3 = 7, R_{k+2} = R_{k-1} + R_k + R_{k+1}$.

By the same method as in this example, one can easily compute the Hausdorff dimension of X_Ω where is an arbitrary (multi-step) shift of finite type.

Example 3.4 (Multiplicative β-shift)**.** *Let $\beta > 1$ and $\Omega = \Omega_\beta$ be the β-shift (see* [11, 3] *for the definition and basic properties of β-shifts). Let $q = 2$. Then*

$$\dim_H(X_{\Omega_\beta}) = \log_2 t, \quad \textit{where} \quad t = \sqrt{d_1 t + \sqrt{d_2 t + \sqrt{d_3 t + \dots}}}, \tag{23}$$

and $d_1 d_2 d_3 \dots$ is the infinite greedy expansion of 1 in base β. Moreover, $\dim_M(X_{\Omega_\beta}) < \dim_H(X_{\Omega_\beta})$ for all $\beta \notin \mathbb{N}$.

Notice that X_{Ω_β} is the multiplicative golden mean shift X_g when $\beta = \frac{1+\sqrt{5}}{2}$, for which the infinite β-expansion of 1 is $101010\dots$

The equation (23) may be justified as follows. Assume that $\beta \notin \mathbb{N}$. By [11], $x = (x_k)_1^\infty \in \Omega_\beta$ if and only if every shift of x, that is $(x_k)_\ell^\infty$ for $\ell \geq 2$, is less than or equal to $(d_k)_1^\infty$ in the lexicographic order. This implies that the tree of followers of the symbols $0, \dots, d_1 - 1$ in Ω_β is isomorphic to the entire $\mathrm{Pref}(\Omega_\beta)$, and we obtain the following equation at the root from (13):

$$t_\varnothing^2 = d_1 t_\varnothing + t_{d_1}.$$

Similarly, we obtain

$$t_{d_1 \dots d_n}^2 = d_{n+1} t_\varnothing + t_{d_1 \dots d_{n+1}}, \quad n \geq 1,$$

which easily reduces to (23).

4. Proof of Proposition 2.5

Recall that for two partitions α and β, the conditional entropy is defined by

$$H_m^\mu(\alpha|\beta) = \sum_{B\in\beta} \Big(- \sum_{A\in\alpha} \mu(A|B) \log_m \mu(A|B) \Big) \mu(B). \tag{24}$$

Proof of Proposition 2.5(i). We have for $k \geq 2$,

$$H_m^\mu(\alpha_k) = H_m^\mu(\alpha_k|\alpha_1) + H_m^\mu(\alpha_1), \tag{25}$$

by the properties of conditional entropy. From (24),

$$H_m^\mu(\alpha_k|\alpha_1) = \sum_{i=0}^{m-1} p_i H_m^{\mu_i}(\alpha_{k-1}(\Omega_i)),$$

where $p_i = \mu[i]$ and $H_m^{\mu_i}(\alpha_{k-1}(\Omega_i))$ is the entropy of the partition of Ω_i, the follower set of i in Ω, into cylinders of length $k-1$, with respect to the measure μ_i, which is the normalized measure induced by μ on Ω_i. Substituting this and (25) into (17) we obtain

$$\begin{aligned} s(\Omega,\mu) &= \frac{q-1}{q} H_m^{\mu}(\alpha_1) + \frac{1}{q}\sum_{i=0}^{m-1} p_i s(\Omega_i,\mu_i) \\ &= \frac{q-1}{q}\Big[H_m^{\mu}(\alpha_1) + \frac{1}{q-1}\sum_{i=0}^{m-1} p_i s(\Omega_i,\mu_i)\Big]. \end{aligned} \tag{26}$$

Now, the measure μ is completely determined by the probability vector $\mathbf{p} = (p_i)_{i=0}^{m-1}$ and the conditional measures μ_i. The optimization problems on Ω_i are independent, so if μ is optimal for Ω, then μ_i is optimal for Ω_i, for all $i \le m$. Thus,

$$s(\Omega) = \max_{\mathbf{p}} \frac{q-1}{q}\Big[H_m^{\mu}(\alpha_1) + \frac{1}{q-1}\sum_{i=0}^{m-1} p_i s(\Omega_i)\Big].$$

Observe that $H_m^{\mu}(\alpha_1) = -\sum_{i=0}^{m-1} p_i \log_m p_i$. It is well-known that

$$\max_{\mathbf{p}} \sum_{i=0}^{m-1} p_i(a_i - \log_m p_i) = \log_m\Big(\sum_{i=0}^{m-1} m^{a_i}\Big),$$

which is achieved if and only if $p_i = m^{a_i}/\sum_{j=0}^{m-1} m^{a_j}$ for $i = 0,\dots,m-1$. We have $a_i = s(\Omega_i)/(q-1)$, which yields the optimal probability vector

$$\mathbf{p} = (p_i)_{i=0}^{m-1},\quad p_i = \frac{t_i}{t_\varnothing^q},\quad \text{where } t_\varnothing := m^{\frac{s(\Omega)}{q-1}},\quad t_i := m^{\frac{s(\Omega_i)}{q-1}},\ i \le m-1,$$

and

$$t_\varnothing^q = \sum_{i=0}^{m-1} t_i.$$

This is the equation (13) at the root of the graph $\Gamma(\Omega)$. However, the problem is analogous at each vertex, so replacing the set Ω with the set of followers of a prefix and repeating the argument, we obtain it for the entire graph. We also get the formulas (21) for the optimal measure μ from the form of the optimal probability vector above. Observe that the solution $\bar{t}$ of the system (13) which we get this way is in the range $[1, m^{1/(q-1)}]$, where we have uniqueness by Lemma 2.1. (Indeed, for any subtree $\Gamma(\Omega_u)$ of the tree $\Gamma(\Omega)$ we have the outdegree bounded by m, and $s(\Omega_u) \le 1$ by (17) and (20), in view of $H_m^{\mu}(\alpha_k) \le k$.)

This concludes the proof of Proposition 2.5(i), including the uniqueness statement. □

Proof of Proposition 2.5(ii). In order to compute $s(\Omega,\mu)$, it is useful to rewrite it in terms of conditional entropies. We have

$$H^\mu_m(\alpha_{k+1}) = H^\mu_m(\alpha_k) + H^\mu_m(\alpha_{k+1}|\alpha_k).$$

Applying this formula repeatedly, we obtain from (17):

$$s(\Omega,\mu) = \sum_{k=1}^\infty \frac{(q-1)^2 H^\mu_m(\alpha_k)}{q^{k+1}} = \Big(\frac{q-1}{q}\Big)\Big[H^\mu_m(\alpha_1) + \sum_{k=1}^\infty \frac{H^\mu_m(\alpha_{k+1}|\alpha_k)}{q^k}\Big]. \tag{27}$$

Observe that

$$\begin{aligned} H^\mu_m(\alpha_1) &= -\sum_{i=0}^{m-1} \frac{t_i}{t^q_\varnothing}\log_m\big(\frac{t_i}{t^q_\varnothing}\big) \\ &= q\log_m t_\varnothing - \sum_{i=0}^{m-1}\frac{t_i}{t^q_\emptyset}\log_m t_i = q\log_m t_\varnothing - \sum_{i=0}^{m-1}\mu[i]\log_m t_i. \end{aligned}$$

Further,

$$\begin{aligned} H^\mu_m(\alpha_{k+1}|\alpha_k) &= \sum_{[u]\in\alpha_k}\mu[u]\Big(-\sum_{j\colon [uj]\in\alpha_{k+1}}\frac{t_{uj}}{t^q_u}\log_m\frac{t_{uj}}{t^q_u}\Big) \\ &= \sum_{[u]\in\alpha_k}\mu[u]\Big(q\log_m t_u - \sum_{j\colon [uj]\in\alpha_{k+1}}\frac{t_{uj}}{t^q_u}\log_m t_{uj}\Big) \\ &= q\sum_{[u]\in\alpha_k}\mu[u]\log_m t_u - \sum_{[v]\in\alpha_{k+1}}\mu[v]\log_m t_v, \end{aligned}$$

in view of $\mu[uj]=\mu[u]\frac{t_{uj}}{t^q_u}$. Now it is clear that the sum in (27) telescopes, and $s(\Omega,\mu) = (q-1)\log_m t_\varnothing$, as desired. □

We point out that Proposition 2.5(i) is not necessary for the proof of Theorem 2.2, only Proposition 2.5(ii) is needed.

5. Proof of the main theorem 2.2

Proof of Proposition 2.3. Fix a probability measure μ on Ω. We are going to demonstrate that for every $\ell\in\mathbb{N}$,

$$\liminf_{n\to\infty}\frac{-\log_m \mathbb{P}_\mu[x_1^n]}{n} \ge (q-1)^2\sum_{k=1}^{\ell}\frac{H^\mu_m(\alpha_k)}{q^{k+1}} \quad \text{for } \mathbb{P}_\mu\text{-a.e. } x, \tag{28}$$

and

(29)
$$\limsup_{n\to\infty} \frac{-\log_m \mathbb{P}_\mu[x_1^n]}{n} \le (q-1)^2 \sum_{k=1}^{\ell} \frac{H_m^\mu(\alpha_k)}{q^{k+1}} + \frac{(\ell+1)\log_m(2m)}{q^\ell} \quad \text{for } \mathbb{P}_\mu\text{-a.e. } x.$$

Then, letting $\ell \to \infty$ will yield $\dim_{\mathrm{loc}}(\mathbb{P}_\mu, x) = s(\Omega,\mu)$ for $\mathbb{P}_\mu$-a.e. x, as desired.

Fix $\ell \in \mathbb{N}$. To verify (28) and (29), we can restrict ourselves to $n = q^\ell r,\ r \in \mathbb{N}$. (Indeed, if $q^\ell r \le n < q^\ell(r+1)$, then

$$\frac{-\log \mathbb{P}_\mu[x_1^n]}{n} \ge \frac{-\log \mathbb{P}_\mu[x_1^{q^\ell r}]}{q^\ell(r+1)} \ge \frac{r}{r+1} \cdot \frac{-\log \mathbb{P}_\mu[x_1^{q^\ell r}]}{q^\ell r},$$

which implies that

$$\liminf_{n\to\infty} \frac{-\log \mathbb{P}_\mu[x_1^n]}{n} = \liminf_{r\to\infty} \frac{-\log \mathbb{P}_\mu[x_1^{q^\ell r}]}{q^\ell r}.$$

The lim sup is dealt with similarly.)

Let

$$\mathcal{G}_n = \mathcal{G}_{q^\ell r} := \{j \le n :\ \exists\, i > n/q^\ell,\ q \nmid i,\ j \in J_i\} \quad \text{and} \quad \mathcal{H}_n := \{j \le n :\ j \notin \mathcal{G}_n\}.$$

Then we have by the definition (11) of the measure $\mathbb{P}_\mu$:

$$\mathbb{P}_\mu[x_1^n] = \mathbb{P}_\mu[x|\mathcal{G}_n] \cdot \mathbb{P}_\mu[x|\mathcal{H}_n] \tag{30}$$

where $[x|\mathcal{G}_n]$ (resp. $[x|\mathcal{H}_n]$) denotes the cylinder set of $y \in X_\Omega$ whose restriction to $\mathcal{G}_n$ (resp. $[x|\mathcal{H}_n]$) coincides with that of x.

First we work with $\mathbb{P}_\mu[x|\mathcal{G}_n]$. In view of (11) we have

$$\mathbb{P}_\mu[x|\mathcal{G}_n] = \prod_{k=1}^{\ell} \prod_{\substack{\frac{n}{q^k} < i \le \frac{n}{q^{k-1}} \\ q \nmid i}} \mu[x_1^n|J_i]. \tag{31}$$

Note that $x_1^n|J_i$ is a word of length k for $i \in (n/q^k, n/q^{k-1}]$, $q \nmid i$, which is a beginning of a sequence in Ω. Thus, $[x_1^n|J_i]$ is an element of the partition α_k. The random variables $x \mapsto -\log_m \mu[x_1^n|J_i]$ are i.i.d for $i \in (n/q^k, n/q^{k-1}]$, $q \nmid i$, and their expectation equals $H_m^\mu(\alpha_k)$, by the definition of entropy. Note that

$$\#\{i \in (n/q^k, n/q^{k-1}] :\ q \nmid i\} = \left(\frac{q-1}{q}\right)\left(\frac{n}{q^k} - \frac{n}{q^{k-1}}\right) = (q-1)^2 \frac{n}{q^{k+1}}. \tag{32}$$

Fixing k,ℓ with $k \le \ell$ and taking $n = q^\ell r$, $r \to \infty$, we get an infinite sequence of i.i.d. random variables. Therefore, by a version of the Law of Large Numbers, we have

(33)

$$\forall\ k \le \ell, \quad \sum_{\substack{\frac{n}{q^k} < i \le \frac{n}{q^{k-1}} \\ q \nmid i}} \frac{-\log_m \mu[x_1^n|J_i]}{(q-1)^2(n/q^{k+1})} \to H_m^\mu(\alpha_k) \ \text{ as } n = q^\ell r \to \infty, \quad \text{for } \mathbb{P}_\mu\text{-a.e. } x.$$

By (31) and (33), for $\mathbb{P}_\mu$-a.e. x,

(34)

$$\frac{-\log_m \mathbb{P}_\mu[x|\mathcal{G}_n]}{n} = \sum_{k=1}^{\ell} \frac{(q-1)^2}{q^{k+1}} \sum_{\substack{\frac{n}{q^k} < i \le \frac{n}{q^{k-1}} \\ q \nmid i}} \frac{-\log_m \mu[x_1^n|J_i]}{(q-1)^2(n/q^{k+1})} \to \sum_{k=1}^{\ell} \frac{(q-1)^2 H_m^\mu(\alpha_k)}{q^{k+1}}\,.$$

Since $\mathbb{P}_\mu[x_1^n] \le \mathbb{P}_\mu[x|\mathcal{G}_n]$, this proves (28). Observe that (28) suffices for the lower bound $\dim_H(X_\Omega) \ge \dim_H(\mathbb{P}_\mu) \ge s(\Omega,\mu)$, so the rest of the proof of this proposition may be skipped if one is only interested in the computation of $\dim_H(X_\Omega)$.

Next we turn to (29), which requires working with $\mathbb{P}_\mu[x|\mathcal{H}_n]$. In view of (32),

$$\begin{aligned} |\mathcal{H}_n| = n - |\mathcal{G}_n| &= n - \sum_{k=1}^{\ell} (q-1)^2 \frac{nk}{q^{k+1}} \\ &= \frac{n}{q^\ell}\Big[(\ell+1) - \frac{\ell}{q}\Big] && (35) \\ &< \frac{(\ell+1)n}{q^\ell} = (\ell+1)r. && (36) \end{aligned}$$

From (35),

$$\sum_{r=1}^{\infty} 2^{-|\mathcal{H}_{q^\ell r}|} < \infty. \tag{37}$$

Define

$$\mathcal{S}(\mathcal{H}_n) := \big\{x \in X_\Omega:\ \mathbb{P}_\mu[x|\mathcal{H}_n] \le (2m)^{-|\mathcal{H}_n|}\big\}.$$

Clearly,

$$\mathbb{P}_\mu(\mathcal{S}(\mathcal{H}_n)) \le 2^{-|\mathcal{H}_n|},$$

since there are at most $m^{|\mathcal{H}_n|}$ cylinder sets $[x|\mathcal{H}_n]$. In view of (37),

$$\mathbb{P}_\mu\Big(\bigcap_{N\ge 1} \bigcup_{r=N}^{\infty} \mathcal{S}(\mathcal{H}_{q^\ell r})\Big) = 0,$$

hence for $\mathbb{P}_\mu$-a.e. $x \in X_\Omega$ there exists $N(x)$ such that $x \notin \mathcal{S}(\mathcal{H}_n)$ for all $n = q^\ell r \geq N(x)$. For such x and $n \geq N(x)$ we have (the last inequality from (36))

$$\frac{-\log_m \mathbb{P}_\mu[x|\mathcal{H}_n]}{n} < \frac{|\mathcal{H}_n| \log_m(2m)}{n} < \frac{(\ell+1)\log_m(2m)}{q^\ell}.$$

Combining this with (34), which also holds $\mathbb{P}_\mu$-a.e., and with (30), yields (29). □

Proof of Proposition 2.4 and the upper bound in Theorem 2.2. Often upper bounds for the Hausdorff dimension are obtained by explicit efficient coverings, which is easier than getting lower bounds. This is not the case here, a feature shared with self-affine carpets from [1, 10]. In fact, we proceed similarly to [10], by exhibiting the "optimal" measure on the set X_Ω to get an upper bound on the Hausdorff dimension. We use the following well-known result; it essentially goes back to Billingsley [2].

Proposition 5.1 (see [4])**.** *Let E be a Borel set in Σ_m and let ν be a finite Borel measure on Σ_m. If*

$$\liminf_{n\to\infty} \frac{-\log_m \nu[x_1^n]}{n} \leq s \quad \textit{for all } x \in E,$$

then $\dim_H(E) \leq s$.

It should be emphasized that the lower pointwise dimension of ν needs to be estimated from above **for all** $x \in E$, unlike in the proof of the lower bound, where the lower estimate for lim inf is required only ν-a.e.

Lemma 5.2. *Let μ be the measure on Ω defined by (21), and let $\mathbb{P}_\mu$ be the corresponding measure on X_Ω, defined by (16). Then for any $x \in X_\Omega$, denoting*

$$a_\ell(x) := \frac{-\log_m \mathbb{P}_\mu[x_1^n]}{n} \quad \textit{for } n = q^\ell,$$

we have

$$\lim_{\ell\to\infty} \frac{a_1(x) + \cdots + a_\ell(x)}{\ell} = (q-1)\log_m t_\varnothing. \tag{38}$$

Thus, $\liminf_{\ell\to\infty} a_\ell(x) \leq (q-1)\log_m t_\varnothing$ *for all* $x \in X_\Omega$.

Once we prove the lemma, we are done with Theorem 2.2, since by Proposition 5.1 we will then get $\dim_H(X_\Omega) \leq (q-1)\log_m t_\varnothing$. Proposition 2.4 then follows by Proposition 2.5(ii).

Proof of Lemma 5.2. Let $n = q^\ell$ and denote

$$\mathbf{x}_i^{(j)} := x_i x_{qi} \cdots x_{q^j i}.$$

We will also write $t(u)$ for t_u in this proof, to make the formulas more readable. Combining (16) with (21) yields

$$
\begin{aligned}
-\log_m \mathbb{P}_\mu[x_1^n] &= -\sum_{k=1}^{\ell+1} \sum_{\substack{\frac{n}{q^k} < i \le \frac{n}{q^{k-1}} \\ q \nmid i}} \left(\log_m \mu[x_i] + \sum_{j=1}^{k-1} \log_m \frac{\mu[\mathbf{x}_i^{(j)}]}{\mu[\mathbf{x}_i^{(j-1)}]}\right) \\
&= -\sum_{k=1}^{\ell+1} \sum_{\substack{\frac{n}{q^k} < i \le \frac{n}{q^{k-1}} \\ q \nmid i}} \left(\log_m \frac{t(x_i)}{t^q(\varnothing)} + \sum_{j=1}^{k-1} \log_m \frac{t(\mathbf{x}_i^{(j)})}{t^q(\mathbf{x}_i^{(j-1)})}\right). \qquad (39)
\end{aligned}
$$

For $\kappa \in \mathbb{N}$ and $x \in \Omega$ denote

$$\gamma_x(\kappa) := \log_m t(\mathbf{x}_i^{(j)}), \quad \text{where } \kappa = q^j i,\ q \nmid i.$$

Then, telescoping the sum $\sum_{j=1}^{k-1}$ in (39) we obtain

$$-\log_m \mathbb{P}_\mu[x_1^n] = n(q-1)\log_m t(\varnothing) + (q-1)\sum_{\kappa=1}^{n/q} \gamma_x(\kappa) - \sum_{\kappa=n/q+1}^{n} \gamma_x(\kappa).$$

(Note that we pick up $q \log_m t(\varnothing)$ from each number in $[1,n]$ that is not divisible by q, for a total of $n(q-1)\log_m t(\varnothing)$.) Denote

$$S_n := \sum_{\kappa=1}^{n} \gamma_x(\kappa);$$

then

$$-\log_m \mathbb{P}_\mu[x_1^n] = n(q-1)\log_m t(\varnothing) + qS_{n/q} - S_n \quad \text{for } n = q^\ell.$$

We have for $n = q^\ell$, $\ell \ge 1$:

$$a_\ell(x) = \frac{-\log_m \mathbb{P}_\mu[x_1^n]}{n} = (q-1)\log_m t(\varnothing) + \frac{S_{n/q}}{n/q} - \frac{S_n}{n}\,.$$

This implies

$$\frac{a_1 + \cdots + a_\ell}{\ell} = (q-1)\log_m t(\varnothing) + \frac{S_1}{\ell} - \frac{S_{q^\ell}}{\ell q^\ell} \to (q-1)\log_m t(\varnothing), \quad \text{as } \ell \to \infty,$$

as desired. $\square$

Proof of Lemma 2.1. We follow the scheme of the proof of [9, Theorem 5.1].

Let V be the set of vertices of the graph and let M be the maximal outdegree. Consider the space of functions $Y := [1, M^{1/(q-1)}]^V$ from V to $[1, M^{1/(q-1)}]$, which is compact in the topology of pointwise convergence, and the transformation $F : Y \to Y$, given by

$$F(y_v) = \Big(\sum_{w:\ vw \in E} y_w \Big)^{1/q}.$$

(It is easy to see that F maps Y into Y.)

Observe that F is **monotone** in the sense that

$$\overline{y}, \overline{z} \in Y,\ \overline{y} \le \overline{z} \implies F(\overline{y}) \le F(\overline{z}),$$

where "$\le$" is the pointwise partial order. Let $\overline{1}$ be the constant 1 function. Then $\overline{1} \le F(\overline{1}) \le F^2(\overline{1}) \le \dots$ By compactness, there is a pointwise limit

$$\overline{t} = \lim_{n\to\infty} F^n(\overline{1}),$$

which is a fixed point of F, hence $\overline{t}$ satisfies the system of equations (13).

It remains to verify uniqueness. Suppose $\overline{t}$ and $\overline{t'}$ are two distinct fixed points of F. Without loss of generality, we can assume that $\overline{t} \not\le \overline{t'}$. Then let

$$\alpha := \inf\{\xi > 1 :\ \overline{t} \le \xi \overline{t'}\}.$$

Clearly $\alpha \le M^{1/(q-1)}$. By continuity we have $\overline{t} \le \alpha\overline{t'}$, and so $1 < \alpha$ by assumption. Now,

$$\overline{t} = F(\overline{t}) \le F(\alpha\overline{t'}) = \alpha^{1/q} F(\overline{t'}) = \alpha^{1/q}\overline{t'},$$

contradicting the definition of α. The proof is complete. □

Proof of the statements on Minkowski dimension in Theorem 2.2. It is well-known that one can use covering by cylinder sets in the definition of lower Minkowski dimension, so we have for $X \subset \Sigma_m$:

$$\underline{\dim}_M(X) = \liminf_{n\to\infty} \frac{\log_m |\mathrm{Pref}_n(X)|}{n} \tag{40}$$

where $|\mathrm{Pref}_n(X)|$ is the number of prefixes over all sequences in X; equivalently, the number of cylinder sets of length n which intersect X. We get the upper Minkowski dimension $\overline{\dim}(X)$ by replacing lim inf with lim sup in (40).

For dimension computations, we can restrict ourselves to n from an arithmetic progression, so we can take $n = q^\ell r$ for a fixed $\ell \in \mathbb{N}$. Recall that $x \in X_\Omega$ if and only if $x|J_i \in \Omega$ for all i such that $q \nmid i$. It follows that $|\mathrm{Pref}_n(X_\Omega)|$ is bounded below by the product of $|\mathrm{Pref}_k(\Omega)|$ for each $i \in (n/q^k, n/q^{k-1}]$, with $q \nmid i$, over $k = 1, \ldots, \ell$. Thus, in view of (32), we have

$$\log_m |\mathrm{Pref}_n(X_\Omega)| \ge (q-1)^2 \sum_{k=1}^{\ell} \frac{n \log_m |\mathrm{Pref}_k(\Omega)|}{q^{k+1}} .$$

On the other hand,

$$\log_m |\mathrm{Pref}_n(X_\Omega)| \le (q-1)^2 \sum_{k=1}^{\ell} \frac{n \log_m |\mathrm{Pref}_k(\Omega)|}{q^{k+1}} + n - \sum_{k=1}^{\ell} k(q-1)^2 \frac{n}{q^{k+1}}$$

by putting arbitrary digits in the remaining places. Dividing by n and letting $n \to \infty$ we obtain

$$\underline{\dim}_M(X_\Omega) \ge (q-1)^2 \sum_{k=1}^{\ell} \frac{\log_m |\mathrm{Pref}_k(\Omega)|}{q^{k+1}}$$

and

$$\overline{\dim}_M(X_\Omega) \le (q-1)^2 \sum_{k=1}^{\ell} \frac{\log_m |\mathrm{Pref}_k(\Omega)|}{q^{k+1}} + (\ell+1)q^{-\ell} - \ell q^{-\ell-1}.$$

Since $\ell \in \mathbb{N}$ is arbitrary, this yields (15).

It remains to verify that $\dim_M(X_\Omega) = \dim_H(X_\Omega)$ if and only if the tree of prefixes $\Gamma(\Omega)$ is spherically symmetric. Compare the formula (15) with (17). Observe that

$$H_m^\mu(\alpha_k) \le \log_m |\mathrm{Pref}_k(\Omega)|,$$

with equality if and only every cylinder set $[u]$, for $u \in \mathrm{Pref}_k(\Omega)$, has equal measure μ. To get $\dim_H(X_\Omega)$, we have μ the optimal measure from (21). It is immediate from the equations (13) that the solution t_u depends only on the length of the prefix u if and only if $\Gamma(\Omega)$ is spherically symmetric. This implies the desired claim. □

6. Concluding remarks

1. The motivation to consider the multiplicative golden mean shift Ξ_g in [5] came from the study of the dimension spectrum of certain multiple ergodic averages. For $\theta \in [0,1]$ let

$$A_\theta = \Big\{(x_k)_1^\infty \in \Sigma_2 : \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} x_k x_{2k} = \theta\Big\}.$$

The authors of [5] ask what is the Hausdorff dimension of A_θ. It is easy to see that $\dim_H(A_0) = \dim_H(\Xi_g)$, and moreover, recently the methods developed in the present paper have been adapted to compute the full dimension spectrum $\theta \mapsto \dim_H(A_\theta)$ [12]. Independently, the dimension of A_θ and other sets of this type has been computed in [6].

2. Not all subsets of Σ_m that are invariant under the action of multiplicative integers are of the form X_Ω considered in this paper. In fact, the sets of the form X_Ω behave rather like full shifts, because they are "composed" of independent copies of the set Ω, albeit in a "staggered" pattern. On the other hand, let

$$X := \{x \in \Sigma_2 : \ x_k x_{2k} x_{3k} = 0 \text{ for all } k\}.$$

Then clearly

$$(x_k)_{k=1}^\infty \in X \ \Rightarrow \ (x_{rk})_{k=1}^\infty \in X \quad \text{for all } r \in \mathbb{N},$$

but our methods are inadequate to compute the dimension of X.

Acknowledgment. We are grateful to Jörg Schmeling for passing to us the question about the Hausdorff dimension of the "multiplicative golden mean shift" Ξ_g.

References

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

[2] P. Billingsley. *Ergodic theory and information.* Wiley, New York, 1965.

[3] K. Dajani and C. Kraaikamp. *Ergodic theory of numbers.* Carus Mathematical Monographs **29**. Mathematical Association of America, Washington, DC, 2002.

[4] K. Falconer. *Techniques in fractal geometry.* John Wiley & Sons, Chichester, 1997.

[5] A. Fan, L. Liao, J. Ma, *Level sets of multiple ergodic averages*, arXiv:1105.3032.

[6] A. Fan, J. Schmeling, M. Wu, *Multifractal analysis of multiple ergodic averages*, Preprint, 2011.

[7] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1** (1967), 1–49.

[8] R. Kenyon, Y. Peres, B. Solomyak. Hausdorff dimension of the multiplicative golden mean shift. *C. R. Acad. Sci. Paris, Ser. I* **349** (2011), 625–628.

[9] R. Lyons. Random walks and percolation on trees. *Ann. Probab.* **18** (1990), 931–958.

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

[11] W. Parry. On the β-expansions of real numbers. *Acta Math. Acad. Sci. Hungar.* **11** (1960), 401–416.

[12] Y. Peres and B. Solomyak. *Dimension spectrum for a nonconventional ergodic average*, Preprint, 2011.

Richard Kenyon, Department of Mathematics, Brown University, Providence, RI 02912

Yuval Peres, Microsoft Research

Boris Solomyak, Box 354350, Department of Mathematics, University of Washington, Seattle WA 98195

E-mail address: solomyak@math.washington.edu