Entropy Rates and Finite-State Dimension

Chris Bourke* John M. Hitchcock† N. V. Vinodchandran‡

Abstract

The effective fractal dimensions at the polynomial-space level and above can all be equivalently defined as the $\mathcal{C}$-*entropy rate* where $\mathcal{C}$ is the class of languages corresponding to the level of effectivization. For example, pspace-dimension is equivalent to the PSPACE-entropy rate.

At lower levels of complexity the equivalence proofs break down. In the polynomial-time case, the P-entropy rate is a lower bound on the p-dimension. Equality seems unlikely, but separating the P-entropy rate from p-dimension would require proving P $\neq$ NP.

We show that at the finite-state level, the opposite of the polynomial-time case happens: the REG-entropy rate is an upper bound on the finite-state dimension. We also use the finite-state genericity of Ambos-Spies and Busse (2003) to separate finite-state dimension from the REG-entropy rate.

However, we point out that a *block-entropy rate* characterization of finite-state dimension follows from the work of Ziv and Lempel (1978) on finite-state compressibility and the compressibility characterization of finite-state dimension by Dai, Lathrop, Lutz, and Mayordomo (2004).

As applications of the REG-entropy rate upper bound and the block-entropy rate characterization, we prove that every regular language has finite-state dimension 0 and that normality is equivalent to finite-state dimension 1.

1 Introduction

The effective fractal dimensions, introduced by Lutz [20, 21] using success sets of *gales*, can be equivalently formulated using growth rates of *martingales* [3] or log-loss of *predictors* [15] at all levels of complexity. At the polynomial-space, computable, and constructive levels of effectivization, each of these dimensions also admits an *entropy rate* characterization using the corresponding language class [13, 14]. More specifically, polynomial-space dimension is equivalent to the PSPACE-entropy rate, computable dimension is the DEC-entropy rate, and constructive dimension is the CE-entropy rate.

At lower levels of complexity the equivalence proofs for dimension and entropy rates break down. All we know in the polynomial-time case is that the P-entropy rate is a lower bound on the p-dimension. Equality seems unlikely, but it follows from recent work [16] that separating the P-entropy rate from p-dimension would require proving P $\neq$ NP.

In this paper we investigate entropy rates at an even lower level of effectivization: finite-state dimension, which was introduced by Dai, Lathrop, Lutz, and Mayordomo [9]. We show in section

*Department of Computer Science and Engineering, University of Nebraska-Lincoln, `cbourke@cse.unl.edu`. Research supported in part by University of Nebraska Layman Award and a NASA-EPSCoR space grant.

†Department of Computer Science, University of Wyoming, `jhitchco@cs.uwyo.edu`.

‡Department of Computer Science and Engineering, University of Nebraska-Lincoln, `vinod@cse.unl.edu`. Research supported in part by NSF grant CCF-0430991 and University of Nebraska Layman Award.

3 that the opposite of the polynomial-time case happens at the finite-state level: the REG-entropy rate is an upper bound on the finite-state dimension. We also observe that the REG-entropy rate behaves more like an effective box-counting dimension than an effective Hausdorff dimension.

In section 4 we establish relationships between the finite-state genericity of Ambos-Spies and Busse [2] and the REG-entropy rate. In particular, an individual sequence is finite-state generic if and only if its REG-entropy rate is 1. By results on the finite-state dimension of frequency classes [9], this immediately implies a separation of finite-state dimension from the REG-entropy rate.

While finite-state dimension is not equivalent to the REG-entropy rate (and it does not seem to admit an entropy rate characterization using any other language class), we point out in section 5 that a *block-entropy rate* characterization of finite-state dimension for individual sequences follows from previous work. Ziv and Lempel [30] showed that the finite-state compressibility of a sequence is equivalent to its block-entropy rate. Combining this with the finite-state compressibility characterization of finite-state dimension [9] yields the equivalence. (In this introduction we are ignoring some asymptotic details involving the difference between dimension and strong dimension [4] that are handled in the body of the paper.) We also develop an extension of this characterization for classes of sequences.

In section 6 we give some applications of the REG-entropy rate upper bound and the block-entropy rate characterization, improving two results from [9]:

(i) Any sequence has finite-state dimension 1 if and only if it is normal.

(ii) Every regular language has finite-state dimension 0.

2 Preliminaries

We write $\{0,1\}^*$ for the set of all finite binary *strings* and $\mathbf{C}$ for the *Cantor space* of all infinite binary *sequences*. A *language* is a subset of $\{0,1\}^*$. In the standard way, a sequence $S \in \mathbf{C}$ can be identified with the language for which it is the characteristic sequence. The length of a string $w \in \{0,1\}^*$ is $|w|$. For a language $A \subseteq \{0,1\}^*$, $A_{=n}$ is the set of all strings in A of length n. The string consisting of the first n bits of $x \in \{0,1\}^* \cup \mathbf{C}$ is denoted by $x \upharpoonright n$ and the substring consisting of the i^{th} through j^{th} bits of x is $x[i..j]$. We write $w \sqsubseteq x$ if w is a prefix of x. For a string $w \in \{0,1\}^*$, $\mathbf{C}_w = \{S \in \mathbf{C} \mid w \sqsubseteq S\}$.

2.1 Finite-State Dimension

Finite-state dimension was developed by Dai, Lathrop, Lutz, and Mayordomo [9] as a generalization of Hausdorff dimension [12]. Later, finite-state strong dimension was similarly introduced by Athreya, Hitchcock, Lutz, and Mayordomo [4] as a generalization of packing dimension [29, 28]. We now recall an equivalent formulation of all these dimensions using log-loss prediction [15, 4].

Definition. A *predictor* is a function $\pi : \{0,1\}^* \times \{0,1\} \to [0,1]$ such that for all $w \in \{0,1\}^*$, $\pi(w,0) + \pi(w,1) = 1$.

Definition. Let π be a predictor, $w \in \{0,1\}^*$, $S \in \mathbf{C}$, and $X \subseteq \mathbf{C}$.

1. The *cumulative log-loss* of π on w is

$$\mathcal{L}^{\log}(\pi, w) = \sum_{i<|w|} \log \frac{1}{\pi(w \upharpoonright i, w[i])}.$$

(We use the convention that $\log \frac{1}{0} = \infty$.)

2. The *log-loss rate* of π on S is

$$\mathcal{L}^{\log}(\pi, S) = \liminf_{n\to\infty} \frac{\mathcal{L}^{\log}(\pi, S\upharpoonright n)}{n}.$$

3. The *worst-case log-loss rate* of π on X is

$$\mathcal{L}^{\log}(\pi, X) = \sup_{S\in X} \mathcal{L}^{\log}(\pi, S).$$

4. The *strong log-loss rate* of π on S is

$$\mathcal{L}^{\log}_{\mathrm{str}}(\pi, S) = \limsup_{n\to\infty} \frac{\mathcal{L}^{\log}(\pi, S\upharpoonright n)}{n}.$$

5. The *worst-case strong log-loss rate* of π on a X is

$$\mathcal{L}^{\log}_{\mathrm{str}}(\pi, X) = \sup_{S\in X} \mathcal{L}^{\log}_{\mathrm{str}}(\pi, S).$$

In [15, 4], the following definitions are shown equivalent to the original definitions of Hausdorff dimension and packing dimension. We refer to [11, 20, 4] for more background on these dimensions.

Definition. Let $X \subseteq \mathbf{C}$. Let Π be the class of all predictors.

1. The *Hausdorff dimension* of X is

$$\dim_{\mathrm{H}}(X) = \inf\{\mathcal{L}^{\log}(\pi, X) \mid \pi \in \Pi\}.$$

2. The *packing dimension* of X is

$$\dim_{\mathrm{P}}(X) = \inf\{\mathcal{L}^{\log}_{\mathrm{str}}(\pi, X) \mid \pi \in \Pi\}.$$

The finite-state dimensions may be similarly defined by using predictors that arise from finite-state gamblers.

Definition. A *finite-state gambler (FSG)* is a tuple $G = (Q, \delta, \beta, q_0)$ where

- Q is a nonempty, finite set of states,
- $\delta : Q \times \{0, 1\} \to Q$ is the transition function,
- $\beta : Q \times \{0, 1\} \to \mathbb{Q} \cap [0, 1]$ is the *betting function*, which satisfies

  $$\beta(q, 0) + \beta(q, 1) = 1$$

  for all $q \in Q$, and
- $q_0 \in Q$ is the initial state.

An FSG $G = (Q, \delta, \beta, q_0)$ defines a predictor π_G by

$$\pi_G(w, a) = \beta(\delta^*(w), a)$$

for all $w \in \{0,1\}^*$ and $a \in \{0,1\}$. Here $\delta^* : \{0,1\}^* \to Q$ is the standard extension of δ to strings defined by the recursion

$$\delta^*(\lambda) = q_0,$$
$$\delta^*(wa) = \delta(\delta^*(w), a).$$

We say that a predictor π is *finite-state* if $\pi = \pi_G$ for some FSG G.

Definition. Let $X \subseteq \mathbf{C}$. Let $\Pi(\text{FS})$ be the class of all finite-state predictors.

1. The *finite-state dimension* of X is

$$\dim_{\text{FS}}(X) = \inf\{\mathcal{L}^{\log}(\pi, X) \mid \pi \in \Pi(\text{FS})\}.$$

2. The *finite-state strong dimension* of X is

$$\text{Dim}_{\text{FS}}(X) = \inf\{\mathcal{L}^{\log}_{\text{str}}(\pi, X) \mid \pi \in \Pi(\text{FS})\}.$$

The following holds for every $X \subseteq \mathbf{C}$:

$$\begin{array}{ccccccc} 0 & \leq & \dim_{\text{H}}(X) & \leq & \dim_{\text{FS}}(X) & & \\ & & \rotatebox{90}{$\geq$} & & \rotatebox{90}{$\geq$} & & \\ & & \dim_{\text{P}}(X) & \leq & \text{Dim}_{\text{FS}}(X) & \leq & 1. \end{array}$$

We will also consider the finite-state dimensions of individual sequences.

Definition. Let $S \in \mathbf{C}$.

1. The *finite-state dimension* of S is $\dim_{\text{FS}}(S) = \dim_{\text{FS}}(\{S\})$.
2. The *finite-state strong dimension* of S is $\text{Dim}_{\text{FS}}(S) = \text{Dim}_{\text{FS}}(\{S\})$.

The following proposition states that changing an initial segment of a sequence does not change its finite-state dimension.

Proposition 2.1. *For all* $S \in \mathbf{C}$ *and* $x, y \in \{0,1\}^*$, $\dim_{\text{FS}}(xS) = \dim_{\text{FS}}(yS)$ *and* $\text{Dim}_{\text{FS}}(xS) = \text{Dim}_{\text{FS}}(yS)$.

2.2 Entropy Rates

We now review entropy rates of languages and their relationship to dimension. The following concept dates back to Chomsky and Miller [7] and Kuich [17].

Definition. Let $A \subseteq \{0,1\}^*$. The *entropy rate of* A is

$$H_A = \limsup_{n \to \infty} \frac{\log |A_{=n}|}{n}.$$

Intuitively, H_A gives an asymptotic measurement of the amount by which every string in $A_{=n}$ is compressed in an optimal code. The following equivalent definition of H_A is useful in some contexts.

Lemma 2.2. (Staiger [26]) *For any* $A \subseteq \{0,1\}^*$,

$$H_A = \inf \left\{ s \,\middle|\, \sum_{w \in A} 2^{-s|w|} < \infty \right\}.$$

For any language A we define two classes of sequences $A^{\mathrm{i.o.}}$ and $A^{\mathrm{a.e.}}$.

Definition. Let $A \subseteq \{0,1\}^*$.

1. The *i.o.-class of* A is $A^{\mathrm{i.o.}} = \{S \in \mathbf{C} \mid (\exists^\infty n) S\upharpoonright n \in A\}$.
2. The *a.e.-class of* A is $A^{\mathrm{a.e.}} = \{S \in \mathbf{C} \mid (\forall^\infty n) S\upharpoonright n \in A\}$.

The name *δ-limit of A* and notation A^δ have also been used for $A^{\mathrm{i.o.}}$ [26, 27].

Definition. Let $\mathcal{C}$ be a class of languages and $X \subseteq \mathbf{C}$.

1. The *$\mathcal{C}$-entropy rate* of X is
$$\mathcal{H}_{\mathcal{C}}(X) = \inf\{H_A \mid A \in \mathcal{C} \text{ and } X \subseteq A^{\mathrm{i.o.}}\}.$$
2. The *strong $\mathcal{C}$-entropy rate* of X is
$$\mathcal{H}^{\mathrm{str}}_{\mathcal{C}}(X) = \inf\{H_A \mid A \in \mathcal{C} \text{ and } X \subseteq A^{\mathrm{a.e.}}\}.$$

Informally, $\mathcal{H}_{\mathcal{C}}(X)$ is the lowest entropy rate with which every element of X can be covered infinitely often by a language in $\mathcal{C}$.

For all $X \subseteq \mathbf{C}$, classical results (see [23, 26]) imply

$$\dim_{\mathrm{H}}(X) = \mathcal{H}_{\mathrm{ALL}}(X),$$

where ALL is the class of all languages and $\dim_{\mathrm{H}}$ is Hausdorff dimension. It is also known [4] that packing dimension is the corresponding strong entropy rate:

$$\dim_{\mathrm{P}}(X) = \mathcal{H}^{\mathrm{str}}_{\mathrm{ALL}}(X).$$

Using other classes of languages gives equivalent definitions of the constructive, computable, and polynomial-space dimensions (see [13, 14, 4, 16] for definitions and more details): for all $X \subseteq \mathbf{C}$,

$$\mathrm{cdim}(X) = \mathcal{H}_{\mathrm{CE}}(X),\ \dim_{\mathrm{comp}}(X) = \mathcal{H}_{\mathrm{DEC}}(X),\ \dim_{\mathrm{pspace}}(X) = \mathcal{H}_{\mathrm{PSPACE}}(X)$$

and

$$\mathrm{cDim}(X) = \mathcal{H}^{\mathrm{str}}_{\mathrm{CE}}(X),\ \mathrm{Dim}_{\mathrm{comp}}(X) = \mathcal{H}^{\mathrm{str}}_{\mathrm{DEC}}(X),\ \mathrm{Dim}_{\mathrm{pspace}}(X) = \mathcal{H}^{\mathrm{str}}_{\mathrm{PSPACE}}(X).$$

In the polynomial-time setting, all that we know is $\mathcal{H}_{\mathrm{P}}(X) \le \dim_{\mathrm{p}}(X)$ and $\mathcal{H}^{\mathrm{str}}_{\mathrm{P}}(X) \le \mathrm{Dim}_{\mathrm{p}}(X)$ always hold.

3 Regular Entropy Rate

In this section we study $\mathcal{H}_{\mathrm{REG}}$, the regular entropy rate, and its relationships with box-counting dimension and finite-state dimension.

3.1 Upper Bound on Box-Counting Dimension

We will show that $\mathcal{H}_{\mathrm{REG}}$ is an upper bound on the box-counting dimension. For any set $X \subseteq \mathbf{C}$ and $n \in \mathbb{N}$, let

$$N_n(X) = |\{S\restriction n \mid S \in X\}|$$

be how many distinct strings of length n are prefixes of elements of X. Then the *(upper) box-counting dimension of* X (see [11]) is

$$\overline{\dim}_{\mathrm{B}}(X) = \limsup_{n\to\infty} \frac{\log N_n(X)}{n}.$$

We will use an everywhere version of the infinitely-often and almost-everywhere classes $A^{\text{i.o.}}$ and $A^{\text{a.e.}}$.

Definition. For any $A \subseteq \{0,1\}^*$, let $A^{\square} = \{S \in \mathbf{C} \mid (\forall n) S\restriction n \in A\}$.

Using $A^{\square}$, we can define a concept similar to the entropy rates.

Definition. For any $X \subseteq \mathbf{C}$ and class $\mathcal{C}$ of languages, let

$$\mathcal{H}^{\square}_{\mathcal{C}}(X) = \inf\{H_A \mid X \subseteq A^{\square} \text{ and } A \in \mathcal{C}\}.$$

When the class of languages is unrestricted in this definition, we get the box-counting dimension.

Proposition 3.1. *For every $X \subseteq \mathbf{C}$, $\overline{\dim}_{\mathrm{B}}(X) = \mathcal{H}^{\square}_{\mathrm{ALL}}(X)$.*

We will see that $\mathcal{H}_{\mathrm{REG}}$ and $\mathcal{H}^{\mathrm{str}}_{\mathrm{REG}}$ are *both* equivalent to $\mathcal{H}^{\square}_{\mathrm{REG}}$. First, we need some notation and a lemma.

Notation. For any $A \subseteq \{0,1\}^*$, let $\mathrm{pref}(A) = \{w \in \{0,1\}^* \mid (\exists x \in A) w \sqsubseteq x\}$.

Lemma 3.2. (Staiger [26]) *For every $A \in \mathrm{REG}$, $H_A = H_{\mathrm{pref}(A)}$.*

Now we can see that the REG-entropy rate behaves like a finite-state box-counting dimension, and that there is no difference between it and the strong REG-entropy rate.

Theorem 3.3. *For every $X \subseteq \mathbf{C}$, $\mathcal{H}_{\mathrm{REG}}(X) = \mathcal{H}^{\mathrm{str}}_{\mathrm{REG}}(X) = \mathcal{H}^{\square}_{\mathrm{REG}}(X)$.*

Proof. The inequalities $\mathcal{H}_{\mathrm{REG}}(X) \leq \mathcal{H}^{\mathrm{str}}_{\mathrm{REG}}(X) \leq \mathcal{H}^{\square}_{\mathrm{REG}}(X)$ are immediate from the definitions. Let $s > \mathcal{H}_{\mathrm{REG}}(X)$. It suffices to show that $\mathcal{H}^{\square}_{\mathrm{REG}}(X) \leq s$. Let $A \in \mathrm{REG}$ such that $H_A < s$ and $X \subseteq A^{\text{i.o.}}$. Then $\mathrm{pref}(A) \in \mathrm{REG}$ and $X \subseteq \mathrm{pref}(A)^{\square}$. By Lemma 3.2 we have $H_{\mathrm{pref}(A)} < s$, so $\mathcal{H}^{\square}_{\mathrm{REG}}(X) \leq s$. $\square$

By Proposition 3.1, it follows that the box dimension is a lower bound on the regular entropy rate.

Corollary 3.4. *For every $X \subseteq \mathbf{C}$, $\overline{\dim}_{\mathrm{B}}(X) \leq \mathcal{H}_{\mathrm{REG}}(X)$.*

3.2 Upper Bound on Finite-State Dimension

Next we show that the REG-entropy rate is always an upper bound on the finite-state strong dimension.

Theorem 3.5. *For any* $X \subseteq \mathbf{C}$, $\mathrm{Dim}_{\mathrm{FS}}(X) \leq \mathcal{H}_{\mathrm{REG}}(X)$.

Proof. If X is empty, then the statement trivially holds, so assume $X \neq \emptyset$. Let $t > s > \mathcal{H}_{\mathrm{REG}}(X) = \mathcal{H}^{\square}_{\mathrm{REG}}(X)$ and let $0 < \epsilon < t - s$. It suffices to show that $\mathrm{Dim}_{\mathrm{FS}}(X) \leq t$. Let $A \in \mathrm{REG}$ such that $X \subseteq A^{\square}$ and $H_A < s$. Since X is not empty, we have $A \neq \emptyset$.

Let $M = (Q, \delta, q_0, F)$ be a minimal DFA for A. For each $q \in Q$, let

$$W_q = \{w \in \{0,1\}^* \mid \delta(q, w) \in F\}$$

and

$$m(q) = \sum_{w \in W_q} 2^{-s|w|}.$$

Since M is a minimal DFA, for each q there is some string x_q such that $\delta(q_0, x_q) = q$. Let

$$A(x_q) = \{w \in A \mid x_q \sqsubseteq w\} = x_q W_q.$$

We have

$$m(q) = 2^{s|x_q|} \sum_{w \in A(x_q)} 2^{-s|w|} \leq 2^{s|x_q|} \sum_{w \in A} 2^{-s|w|},$$

which is finite by Lemma 2.2. Note that for any $q \in Q$, we have

$$0W_{\delta(q,0)} \cup 1W_{\delta(q,1)} \subseteq W_q,$$

so

$$m(\delta(q,0)) + m(\delta(q,1)) \leq 2^s m(q).$$

Define a betting-function function $\beta : Q \times \{0,1\} \to [0,1]$ by

$$\beta(q, b) = \frac{m(\delta(q,b))}{m(\delta(q,0)) + m(\delta(q,1))}$$

if the denominator is not 0, and $\beta(q,b) = \frac{1}{2}$ otherwise. Since β may not be rational-valued, let $\hat{\beta} : Q \times \{0,1\} \to [0,1] \cap \mathbb{Q}$ be a betting function approximating β in the sense that for all q and b, $|\log \beta(q,b) - \log \hat{\beta}(q,b)| < \epsilon$. Let G be the finite-state gambler $G = (Q, \delta, \hat{\beta}, q_0)$, and let π_G be the finite-state predictor associated with G.

Let $w \in A$. For each i $(0 \leq i \leq |w|)$, let $q_i = \delta(q_0, w \upharpoonright i)$. We have

$$
\begin{aligned}
\mathcal{L}^{\log}(\pi_G, w) &= \sum_{i=0}^{|w|-1} -\log \pi_G(w \upharpoonright i, w[i]) \\
&= \sum_{i=0}^{|w|-1} -\log \hat{\beta}(q_i, w[i]) \\
&\leq \epsilon|w| + \sum_{i=0}^{|w|-1} -\log \beta(q_i, w[i]) \\
&= \epsilon|w| + \log \prod_{i=0}^{|w|-1} \frac{m(\delta(q_i, 0)) + m(\delta(q_i, 1))}{m(q_{i+1})} \\
&\leq \epsilon|w| + \log \prod_{i=0}^{|w|-1} \frac{2^s m(q_i)}{m(q_{i+1})} \\
&= (s + \epsilon)|w| + \log \frac{m(q_0)}{m(q_{|w|})}.
\end{aligned}
$$

(The assumption $w \in A$ is important here because it implies $m(q_i)$ is always nonzero.) It follows that $\mathcal{L}^{\log}_{\text{str}}(\pi_G, S) \leq t$ for any $S \in A^{\square}$. Therefore $\mathcal{L}^{\log}_{\text{str}}(\pi_G, X) \leq t$, so $\text{Dim}_{\text{FS}}(X) \leq t$. $\square$

4 Finite-State Genericity

This section establishes some connections between regular entropy rates and the finite-state genericity of Ambos-Spies and Busse [2]. From this we will see a separation of the regular entropy rate from finite-state dimension. We first recall the concepts we need from [2]. A function $f : \{0,1\}^* \to \{0,1\}^*$ is *finite-state computable* if there is a DFA M along with strings labeling each of the states such that $f(w)$ is always the label for the state M is in after processing w.

Definition. Let $S \in \mathbf{C}$.

1. S *meets* a function $f : \{0,1\}^* \to \{0,1\}^*$ if for some n we have

$$(S \upharpoonright n) f(S \upharpoonright n) \sqsubseteq S.$$

2. S is *finite-state generic* if S meets every finite-state $f : \{0,1\}^* \to \{0,1\}^*$.

Ambos-Spies and Busse prove that several other definitions are equivalent to this definition of finite-state genericity.

Recall that a set $X \subseteq \mathbf{C}$ is *nowhere dense* if it is contained in the complement of a dense, open set. Equivalently, X is nowhere dense if

$$(\forall w)(\exists w' \sqsupseteq w) X \cap \mathbf{C}_{w'} = \emptyset.$$

In intuitive terms, X is full of holes: given any string w, we can always find an extension w' that is not a prefix of any sequence in X. We now define an effective version of nowhere density where a finite-state function can always identify one of these holes.

Definition. We say that X is *finite-state nowhere dense* if there is a finite-state function $f : \{0,1\}^* \to \{0,1\}^*$ such that
$$(\forall w) X \cap \mathbf{C}_{wf(w)} = \emptyset.$$

This concept leads to another definition of finite-state genericity.

Proposition 4.1. *A sequence $S \in \mathbf{C}$ is finite-state generic if and only if S is not contained in any finite-state nowhere dense set.*

Proof. Assume that S is not finite-state generic. Let f be a finite-state function which S does not meet. Then $X_f = \{T \in \mathbf{C} \mid T \text{ does not meet } f\}$ is finite-state nowhere dense (via f) and contains S.

Now assume that S is contained in some finite-state nowhere dense set X. Let f be a finite-state function showing that X is finite-state nowhere dense. Then S does not meet f, so S is not finite-state generic. □

4.1 Entropy Rates and Genericity

Notation. For any $A \subseteq \{0,1\}^*$ and $x \in \{0,1\}^*$, let
$$A_x = \{w \in A \mid x \sqsubseteq w\}$$
be the set of all extensions of x in A.

The following lemma is essentially a restatement of Lemma 3.2.

Lemma 4.2. *Let $A \in \mathrm{REG}$ and suppose that for infinitely many n,*
$$|\{x \in \{0,1\}^n \mid A_x \neq \emptyset\}| \geq 2^{sn}.$$
Then $H_A \geq s$.

Proof. Recall from Lemma 3.2 that $H_A = H_{\mathrm{pref}(A)}$. If $A_x \neq \emptyset$, then $x \in \mathrm{pref}(A)$, so the hypothesis says $|\mathrm{pref}(A)_{=n}| \geq 2^{sn}$ for infinitely many n. Therefore $H_{\mathrm{pref}(A)} \geq s$. □

We now show a relationship between the regular entropy rate and finite-state nowhere dense sets.

Theorem 4.3. *For every $X \subseteq \mathbf{C}$, $\mathcal{H}_{\mathrm{REG}}(X) < 1$ if and only if X is finite-state nowhere dense.*

Proof. Assume that $\mathcal{H}_{\mathrm{REG}}(X) < s < 1$. Then there is an $A \in \mathrm{REG}$ with $H_A < s$ and $X \subseteq A^{\text{i.o.}}$. By Lemma 4.2 we know that for some n_0, for all $n \geq n_0$,
$$|\{x \in \{0,1\}^n \mid A_x \neq \emptyset\}| < 2^{sn}. \tag{4.1}$$

Let $M = (Q, \delta, q_0, F)$ be the minimal DFA that decides A. For each $q \in Q$, let w_q be a string of minimal length with $\delta^*(q_0, w_q) = q$. Define
$$w'_q = \begin{cases} w_q & \text{if } |w_q| \geq n_0 \\ w_q 0^{n_0 - |w_q|} & \text{otherwise.} \end{cases}$$

Let l be large enough so that $2^{s(|w'_q|+l)} < 2^l$ for all $q \in Q$. Then by (4.1), for each $q \in Q$ there is some $x_q \in \{0,1\}^l$ with $A_{w'_q x_q} = \emptyset$. In each state q, our finite-state function outputs x_q if $|w_q| \geq n_0$, $0^{n_0 - |w_q|} x_q$ if $|w_q| < n_0$. This function shows that X is finite-state nowhere dense.

For the other direction, assume that X is finite-state nowhere dense, and let f be a finite-state function witnessing this. We can assume that $f : \{0,1\}^* \to \{0,1\}^k$ for some $k > 0$. Let

$$A = \{x \mid (\forall m < |x|/k)\ (x \upharpoonright mk) f(x \upharpoonright mk) \not\sqsubseteq x\}.$$

Then $X \subseteq A^{\text{i.o.}}$ and A is regular, so $\mathcal{H}_{\text{REG}}(X) \leq H_A$. Now we will verify that $H_A < 1$. Let n be any length and write $n = mk + l$ where $l < k$. An upper bound on $|A_{=n}|$ is $(2^k - 1)^m \cdot 2^l$, so

$$\frac{\log |A_{=n}|}{n} \leq \frac{l + m \log(2^k - 1)}{n} \leq \frac{k}{n} + \frac{\log(2^k - 1)}{k}$$

and we obtain

$$H_A \leq \frac{\log(2^k - 1)}{k} < 1.$$

□

Combining Theorem 4.3 with Proposition 4.1, we obtain the following corollaries. We write $\mathcal{H}_{\text{REG}}(S) = \mathcal{H}_{\text{REG}}(\{S\})$ for any sequence $S \in \mathbf{C}$.

Corollary 4.4. *A sequence $S \in \mathbf{C}$ is finite-state generic if and only if $\mathcal{H}_{\text{REG}}(S) = 1$.*

Corollary 4.5. *If a set $X \subseteq \mathbf{C}$ contains a finite-state generic sequence, then $\mathcal{H}_{\text{REG}}(X) = 1$.*

A sequence $S \in \mathbf{C}$ is *saturated* if it contains every finite binary string as a substring. Ambos-Spies and Busse [2] showed a sequence is finite-state generic if and only if it is saturated. Therefore Corollary 4.4 can be restated as follows.

Corollary 4.6. *For every $S \in \mathbf{C}$, $\mathcal{H}_{\text{REG}}(S) = 1$ if and only if S is saturated.*

4.2 Separation of Dimension from Entropy Rates

We now separate the regular entropy rate from finite-state strong dimension. Recall from [9] that the class

$$\text{FREQ}_\alpha = \left\{ S \in \mathbf{C} \,\middle|\, \lim_{n \to \infty} \frac{\#(1, S \upharpoonright n)}{n} = \alpha \right\}$$

has finite-state dimension

$$\dim_{\text{FS}}(\text{FREQ}_\alpha) = \mathcal{H}(\alpha) = \alpha \log \frac{1}{\alpha} + (1 - \alpha) \log \frac{1}{1 - \alpha}$$

for every $\alpha \in [0,1]$. In fact, the proof also shows that $\text{Dim}_{\text{FS}}(\text{FREQ}_\alpha) = \mathcal{H}(\alpha)$. Since FREQ_α is dense for all α, we obtain

$$\mathcal{H}_{\text{REG}}(\text{FREQ}_\alpha) = 1$$

from Theorem 4.3. Therefore (using $\alpha \neq \frac{1}{2}$) we see that proper inequality can hold in Theorem 3.5.

In fact, the we can get the same separation for singletons. If we take a sequence $S \in \text{FREQ}_\alpha$ that is saturated, then $\mathcal{H}_{\text{REG}}(S) = 1$ by Corollary 4.6 but $\text{Dim}_{\text{FS}}(S) \leq \mathcal{H}(\alpha)$.

5 Block-Entropy Rate

In this section we use a more general entropy notion, the block-entropy rate, to characterize the finite-state dimensions. This is interesting because the block-entropy rate considers only frequency properties of the sequence and does not involve finite-state machines.

5.1 Finite-State Dimension and Compressibility

First we recall the relationships between finite-state dimension and finite-state compressibility [9, 4].

Definition. A *finite-state compressor (FSC)* is a tuple $C = (Q, \delta, \nu, q_0)$, where

- Q is a nonempty, finite set of states,
- $\delta : Q \times \{0, 1\} \to Q$ is the transition function,
- $\nu : Q \times 0, 1 \to \{0, 1\}^*$ is the output function, and
- $q_0 \in Q$ is the initial state.

The *output* of C on an input $w \in \{0, 1\}^*$ is the string $C(w)$ defined by the recursion

$$\begin{aligned} C(\lambda) &= \lambda, \\ C(xb) &= C(x)\nu(\delta^*(x), b), \end{aligned}$$

for all $x \in \{0, 1\}^*$ and $b \in \{0, 1\}$, where δ^* is defined as in Section 2. We say that C is *information-lossless* if the function $w \mapsto (C(w), \delta^*(w))$ is one-to-one.

Let $\mathcal{C}$ be the collection of all information-lossless finite-state compressors. For each $k \in N$, let $\mathcal{C}_k$ be the collection of all k-state information-lossless finite-state compressors. For any $S \in \mathbf{C}$, define

$$\rho_{\mathrm{FS}}(S) = \inf_{C \in \mathcal{C}} \liminf_{n \to \infty} \frac{|C(S \upharpoonright n)|}{n}$$

and

$$R_{\mathrm{FS}}(S) = \inf_{k \in \mathbb{N}} \limsup_{n \to \infty} \min_{C \in \mathcal{C}_k} \frac{|C(S \upharpoonright n)|}{n}.$$

The quantity $R_{\mathrm{FS}}(S)$ was originally called $\rho(S)$ in [30]. In [9], $\rho(S)$ was modified to obtain $\rho_{\mathrm{FS}}(S)$ and a compressibility characterization of finite-state dimension.

Theorem 5.1. (Dai, Lathrop, Lutz, and Mayordomo [9]) *For every $S \in \mathbf{C}$,*

$$\dim_{\mathrm{FS}}(S) = \rho_{\mathrm{FS}}(S).$$

Later, when strong dimension was introduced, it was shown that $R_{\mathrm{FS}}(S)$ characterizes finite-state strong dimension.

Theorem 5.2. (Athreya, Hitchcock, Lutz, and Mayordomo [4]) *For every $S \in \mathbf{C}$,*

$$\mathrm{Dim}_{\mathrm{FS}}(S) = R_{\mathrm{FS}}(S).$$

5.2 Block Entropy and Compressibility

Let $n, l \in \mathbb{N}$ where l divides n. Given a string $x \in \{0,1\}^n$ and a string $w \in \{0,1\}^l$, let

$$N(w,x) = |\{0 \le i < n/l \mid x[il..(i+1)l-1] = w\}|$$

be the number of times w occurs in the length-l blocks of x. The *relative frequency of w in x* is

$$P(w,x) = \frac{l}{n} N(w,x).$$

The l^{th} *block entropy of x* is

$$H_l(x) = \frac{1}{l} \sum_{w \in \{0,1\}^l} P(w,x) \log \frac{1}{P(w,x)},$$

i.e., the normalized entropy of the distribution $P(\cdot, x)$ on $\{0,1\}^l$.

Definition. Let $S \in \mathbf{C}$.

1. The l^{th} *block-entropy rate of S* is

$$H_l(S) = \liminf_{k \to \infty} H_l(S \upharpoonright kl).$$

2. The *block-entropy rate of S* is

$$H(S) = \inf_{l \in \mathbb{N}} H_l(S).$$

3. The l^{th} *upper block-entropy rate of S* is

$$\overline{H}_l(S) = \limsup_{k \to \infty} H_l(S \upharpoonright kl).$$

4. The *upper block-entropy rate of S* is

$$\overline{H}(S) = \inf_{l \in \mathbb{N}} \overline{H}_l(S).$$

Ziv and Lempel showed that the upper block-entropy rate corresponds to the finite-state compressibility of a sequence.

Theorem 5.3. (Ziv and Lempel [30]) *For every $S \in \mathbf{C}$, $R_{\mathrm{FS}}(S) = \overline{H}(S)$.*

5.3 Block Entropy and Dimension

From Theorems 5.2 and 5.3, we have the following block-entropy rate characterization of finite-state strong dimension.

Theorem 5.4. *For every $S \in \mathbf{C}$, $\mathrm{Dim}_{\mathrm{FS}}(S) = \overline{H}(S)$.*

Does the analogous characterization $\dim_{\rm FS}(S) = H(S)$ hold for finite-state dimension? We will show that it does, establishing it as a corollary of a more general characterization theorem for classes of sequences.

For any $S \in \mathbf{C}$ and compressor $C \in \mathcal{C}$, let

$$\rho_C(S) = \liminf_{n\to\infty} \frac{|C(S\restriction n)|}{n}$$

and let $\overline{\rho_C}(S)$ be the corresponding lim sup. From the proofs of Theorems 5.1 and 5.2 in [9, 4] for individual sequences, it is straightforward to see the following for classes.

Theorem 5.5. *For every $X \subseteq \mathbf{C}$,*

$$\dim_{\rm FS}(X) = \inf_{C\in\mathcal{C}} \sup_{S\in X} \rho_C(S)$$

and

$$\mathrm{Dim}_{\rm FS}(X) = \inf_{C\in\mathcal{C}} \sup_{S\in X} \overline{\rho_C}(S).$$

We will also need the following three lemmas.

Lemma 5.6. *Let $l \in \mathbb{N}$. There exists a compressor $C_l \in \mathcal{C}$ such that for all $S \in \mathbf{C}$, $\rho_{C_l}(S) \leq H_l(S) + 2/l$ and $\overline{\rho_{C_l}}(S) \leq \overline{H_l}(S) + 2/l$.*

Proof. Fix $l \in \mathbb{N}$. From Sheinwald's proof [25] of Theorem 5.3 we know that for every $x \in \{0,1\}^*$ there is a compressor $C_x \in \mathcal{C}_{2^l}$ (using Huffman coding) such that

$$\frac{|C_x(x)|}{|x|} \leq H_l(x) + \frac{1}{l}.$$

From the proof of Theorem 5.2 given in [4], we obtain a compressor C_l such that for all $C \in \mathcal{C}_{2^l}$ and $x \in \{0,1\}^*$,

$$|C_l(x)| \leq |C(x)| + \frac{|x|}{l} + c_l,$$

where c_l is a constant. Therefore for all x,

$$\frac{|C_l(x)|}{|x|} \leq H_l(x) + \frac{2}{l} + \frac{c_l}{|x|},$$

so we have $\rho_{C_l}(S) \leq H_l(S) + 2/l$ for all $S \in \mathbf{C}$. The proof of the second inequality is analogous. □

Lemma 5.7. *Let $C \in \mathcal{C}$ be a compressor. There is a constant c such that for all $l \in \mathbb{N}$ and $S \in \mathbf{C}$, $H_l(S) \leq \rho_C(S) + (c + \log l)/l$ and $\overline{H_l}(S) \leq \overline{\rho_C}(S) + (c + \log l)/l$.*

Proof. Let σ be the number of states in C and let r_C be the maximum number of bits that C outputs on a single transition. From Sheinwald's proof [25] of Theorem 5.3, we have

$$\overline{H_l}(S) \leq \overline{\rho_C}(S) + \frac{\log(\sigma^2(1 + lr_c))}{l}$$

for all $S \in \mathbf{C}$ and $l \in \mathbb{N}$. Letting c be a constant such that $c + \log l \geq \log(\sigma^2(1 + lr_C))$ establishes the second inequality. The proof of the first inequality is analogous. □

Lemma 5.8. *Let $S \in \mathbf{C}$. For all $k, l \geq 1$, $\overline{H_{kl}}(S) \leq \overline{H_l}(S)$ and $H_{kl}(S) \leq H_l(S)$.*

Proof. Ziv and Lempel [30] proved that the limit $\lim_{n\to\infty} \overline{H_l}(S)$ exists for all $S \in \mathbf{C}$. From this proof we can extract the inequality

$$(l+m)H_{l+m}(x) \leq lH_l(x) + mH_m(x)$$

for all $x \in \{0,1\}^*$ and $l, m \geq 1$. It follows by induction that for all $k \geq 1$,

$$klH_{kl}(x) \leq klH_l(x),$$

i.e., $H_{kl}(x) \leq H_l(x)$. From this $\overline{H_{kl}}(S) \leq \overline{H_l}(S)$ follows immediately.

To show $H_{kl}(S) \leq H_l(S)$, let $s > H_l(S)$. Then there is an infinite set $J \subseteq \mathbb{N}$ such that for all $j \in J$, $H_l(S \upharpoonright jl) < s$. Fix k. For each $j \in J$, let j' be a multiple of k such that $j \leq j' < j + k$. Then as j becomes large, $|H_l(S \upharpoonright j'l) - H_l(S \upharpoonright jl)| \to 0$. For each $j \in J$, $H_{kl}(S \upharpoonright j'l) \leq H_l(S \upharpoonright j'l)$ from the previous paragraph, so it follows that $H_{kl}(S) < s$. This holds for all $s > H_l(S)$, so $H_{kl}(S) \leq H_l(S)$. □

We now give block-entropy rate characterizations of finite-state dimension and finite-state strong dimension for classes of sequences.

Theorem 5.9. *For every $X \subseteq \mathbf{C}$,*

$$\dim_{\mathrm{FS}}(X) = \inf_{l\in\mathbb{N}} \sup_{S\in X} H_l(S)$$

and

$$\mathrm{Dim}_{\mathrm{FS}}(X) = \inf_{l\in\mathbb{N}} \sup_{S\in X} \overline{H_l}(S).$$

Proof. We prove the finite-state dimension characterization; the argument for strong dimension is analogous.

Let $s > \dim_{\mathrm{FS}}(X)$. Then by Theorem 5.5 there is a compressor $C \in \mathcal{C}$ such that for all $S \in X$, $\rho_C(S) < s$. From Lemma 5.7 we have a constant c such that $H_l(S) \leq s + (c + \log l)/l$ for all $S \in X$ and $l \in \mathbb{N}$. Taking the infimum over all l, we have that the right-hand side is at most s. This holds for all $s > \dim_{\mathrm{FS}}(X)$, so the $\geq$ inequality holds.

Now let s be greater than the right-hand side. Then there is an $l \in \mathbb{N}$ such that $H_l(S) < s$ for all $S \in X$. From Lemma 5.8, we have $H_{kl}(S) \leq H_l(S)$ for all S. Therefore from Lemma 5.6 we obtain for each k a compressor C_{kl} such that $\rho_{C_{kl}}(S) \leq s + 2/kl$ for all $S \in X$. Taking the infimum over all k, we obtain $\dim_{\mathrm{FS}}(X) \leq s$ by Theorem 5.5. □

The dual of Theorem 5.4 follows immediately from Theorem 5.9.

Theorem 5.10. *For every $S \in \mathbf{C}$, $\dim_{\mathrm{FS}}(S) = H(S)$.*

6 Applications

In this section we apply the upper bound of Theorem 3.5 and the equivalence of Theorem 5.10 to prove a few finite-state dimension results.

6.1 Normality

Definition. (Borel [6]) A sequence $S \in \mathbf{C}$ is *normal* if for every $w \in \{0,1\}^*$,

$$\lim_{n\to\infty} \frac{1}{n} \left| \left\{ i < n \middle| S[i..i+|w|-1] = w \right\} \right| = 2^{-|w|}. \tag{6.1}$$

Dai, Lathrop, Lutz, and Mayordomo [9] used the work of Schnorr and Stimm [24] to show that every normal sequence has finite-state dimension 1. We now use the block-entropy rate characterization to prove the converse, yielding that finite-state dimension 1 is equivalent to normality.[1] This result is analogous to Corollary 4.6 that equates saturation with REG-entropy rate 1.

Theorem 6.1. *For every* $S \in \mathbf{C}$*,* $\dim_{\mathrm{FS}}(S) = 1$ *if and only if* S *is normal.*

Proof. As mentioned above, we already know that S is normal implies $\dim_{\mathrm{FS}}(S) = 1$ from [9]. Now assume that S is not normal. We will use Theorem 5.10 to show that $\dim_{\mathrm{FS}}(S) < 1$.

Since S is not normal, there is some string w such that (6.1) fails. Let $l = |w|$. For each i, write $x_i = S[i..i+l-1]$. Then for some $\epsilon > 0$,

$$(\exists^\infty n) \left| \frac{|\{i < n \mid x_i = w\}|}{n} - 2^{-|w|} \right| > \epsilon.$$

This implies that

$$(\exists m < l)(\exists^\infty k) \left| \frac{|\{j < k \mid x_{jl+m} = w\}|}{k} - 2^{-|w|} \right| > \frac{\epsilon}{l}.$$

Fix an m that satisfies the previous line. Obtain a sequence S' from S by removing the first m bits from S. Then

$$(\exists^\infty k) \left| P(w, S' \restriction kl) - 2^{-|w|} \right| > \frac{\epsilon}{l}.$$

Whenever k satisfies the previous line, $P(\cdot, S' \restriction kl)$ is not uniform, so

$$(\exists^\infty k) H_l(S' \restriction kl) < \delta$$

for some fixed $\delta < 1$. Therefore $H_l(S') < \delta$ and we have

$$\dim_{\mathrm{FS}}(S) = \dim_{\mathrm{FS}}(S') = H(S') \le H_l(S') < 1$$

by Proposition 2.1 and Theorem 5.10. □

6.2 Regular Languages

A sequence $S \in \mathbf{C}$ is *rational* if there exist $u, v \in \{0,1\}^*$ such that $S = uv^\infty$. Let $\mathbf{Q}$ be the set of all rational sequences.

Theorem 6.2. (Dai, Lathrop, Lutz, and Mayordomo [9]) $\dim_{\mathrm{FS}}(\mathbf{Q}) = 1$.

> *Remark.* We can use Theorem 5.9 to give an easy proof of Theorem 6.2. Let $l \ge 1$. Define a long string x by concatenating all 2^l strings of length l together. Let $S = x^\infty$. Then $S \in \mathbf{Q}$ and we have $H_l(S) = 1$ since the frequency distribution for blocks of length l is nearly uniform for long prefixes of S. (It is exactly uniform at lengths that are multiples of $|x|$.) We can do this for every l, so $\dim_{\mathrm{FS}}(\mathbf{Q}) = 1$ by Theorem 5.9.

[1] An anonymous referee pointed out that this converse can also be proved using [24].

Since every rational sequence is the characteristic sequence of a regular language [2], Theorem 6.2 implies the following.

Theorem 6.3. $\dim_{\mathrm{FS}}(\mathrm{REG}) = 1$.

In contrast, it is also shown in [9] that $\dim_{\mathrm{FS}}(S) = 0$ for every *individual* $S \in \mathbf{Q}$. We will strengthen this in Theorem 6.7, showing the same for each individual regular language.

The *factor set* $F_l(S)$ of a sequence $S \in \mathbf{C}$ is the set of all finite strings of length l that appear in S. The *factor complexity function* counts the number of factors for each l:

$$p_S(l) = |F_l(S)|.$$

We define an analog of entropy in terms of a sequence's factors:

$$h(S) = \lim_{l \to \infty} \frac{\log p_S(l)}{l}.$$

This gives an upper bound on the regular entropy rate.

Lemma 6.4. *For every $S \in \mathbf{C}$, $\mathcal{H}_{\mathrm{REG}}(S) \leq h(S)$.*

Proof. Let $l \geq 1$ and let $A_l = F_l(S)^*$. Then A_l is regular and $S \in A_l^{\mathrm{i.o.}}$, so

$$\mathcal{H}_{\mathrm{REG}}(S) \leq H_{A_l} = \frac{\log p_S(l)}{l}.$$

This holds for all l, so $\mathcal{H}_{\mathrm{REG}}(S) \leq h(S)$. □

Corollary 6.5. *For any $S \in \mathbf{C}$ with $p_S(l) = 2^{o(l)}$, $\dim_{\mathrm{FS}}(S) = \mathcal{H}_{\mathrm{REG}}(S) = 0$.*

Though "most" sequences are saturated, many well studied sequences satisfy the condition of Corollary 6.5. Specifically, this result gives a new proof that for any $S \in \mathbf{Q}$, $\dim_{\mathrm{FS}}(S) = 0$. Sturmian sequences (see [5]), $S \in \mathbf{C}$ that satisfy $p_S(l) = l + 1$ for all l, also have finite-state dimension 0. Morphic sequences, sequences defined by an iteratively applied mapping $\{0,1\} \mapsto \{0,1\}^*$ have dimension zero since their factor complexity function is quadratic [10].

Automatic sequences are sequences, $(a_n)_{n \geq 0}$ defined by a finite-state function, $f : [n]_k \mapsto \Delta$ where Δ is some finite output alphabet that is applied to each final state. Given the limited computation power of such a model, its not surprising that k-automatic sequences are not too complex.

Theorem 6.6. (Cobham [8]) *For every automatic sequence S, $p_S(l) = O(l)$. In particular, $h(S) = 0$.*

More precisely, $(a_n)_{n \geq 0}$ is defined by feeding a DFA with the canonical representation of n in base-k. For our purposes, we only consider 2-automatic sequences with the same output alphabet $\Delta = \{0,1\}$. In addition, we can equivalently consider $(s_n)_{n \geq 0}$ where s_n is the n^{th} string in the standard enumeration since there exists a finite-state function $g : [n]_2 \mapsto s_n$ (add 1 to $[n]_2$ and drop the leading bit—this can be computed by a simple finite-state transducer). An output mapping of 1 for any $s_n \in L$ and 0 otherwise defines the characteristic sequence of a regular language. For a generalization to any enumeration system see [22].

We now have the result promised earlier: regular languages have finite-state dimension 0.

Theorem 6.7. *For every $A \in \mathrm{REG}$, $\dim_{\mathrm{FS}}(A) = \mathcal{H}_{\mathrm{REG}}(A) = 0$.*

6.3 Morphic Sequences

Automatic sequences are closely related to morphic sequences. A function $\varphi : \{0,1\}^* \to \{0,1\}^*$ is called a *morphism* if $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in \{0,1\}^*$. The iterative application of a morphism φ is defined as $\varphi^0(b) = b$ and $\varphi^i(b) = \varphi(\varphi^{i-1}(b))$ for $b \in \{0,1\}$. A morphism is *expanding* if $|\varphi(b)| \geq 2$ for all $b \in \{0,1\}$. We call a morphism k-uniform if $|\varphi(b)| = k$ for all $b \in \{0,1\}$. A 1-uniform morphism is called a *coding*. Morphisms can be very naturally applied to sequences $S \in \mathbf{C}$,

$$\varphi(S) = \varphi(S[0])\varphi(S[1])\varphi(S[2])\ldots$$

If $\varphi(S) = S$ then φ is called a *fixed point morphism*.

The continued application of an expanding morphism may define a sequence $S \in \mathbf{C}$. If for some $b \in \{0,1\}$ and $x \in \{0,1\}^+$, $\varphi(b) = bx$ then we say that φ is *prolongable* on b. The sequence defined by such a morphism *converges* to

$$S = \varphi^\omega(b) = bx\varphi(x)\varphi^2(x)\varphi^3(x)\ldots$$

which is also a fixed point of φ. That is, $\varphi(\varphi^\omega(b)) = \varphi^\omega(b)$. Such a sequence is called a *pure morphic sequence*. If there is a coding $\tau : \{0,1\} \to \{0,1\}$ such that $S = \tau(\varphi^\omega(b))$ then it is simply a *morphic sequence*.

Theorem 6.8. (Ehrenfeucht and Rozenberg [10]) *The complexity of a sequence $S \in \mathbf{C}$ that is a fixed point of any morphism (not necessarily of constant length) satisfies $p_S(l) \in \mathcal{O}(l^2)$*

Corollary 6.9. *Let $S \in \mathbf{C}$ be a morphic sequence. Then* $\dim_{\mathrm{FS}}(S) = \mathcal{H}_{\mathrm{REG}}(S) = 0$.

Acknowledgments. We thank Peter Bro Miltersen, Jack Lutz, Elvira Mayordomo, and Pascal Weil for helpful comments and discussions. We also thank an anonymous referee for informing us that [24] yields another proof of Theorem 6.1.

References

[1] J.-P. Allouche and J. Shallit. *Automatic Sequences*. Cambridge University Press, first edition, 2003.

[2] K. Ambos-Spies and E. Busse. Automatic forcing and genericity: On the diagonalization strength of finite automata. In *Proceedings of the 4th International Conference on Discrete Mathematics and Theoretical Computer Science*, pages 97–108. Springer-Verlag, 2003.

[3] K. Ambos-Spies, W. Merkle, J. Reimann, and F. Stephan. Hausdorff dimension in exponential time. In *Proceedings of the 16th IEEE Conference on Computational Complexity*, pages 210–217. IEEE Computer Society, 2001.

[4] K. B. Athreya, J. M. Hitchcock, J. H. Lutz, and E. Mayordomo. Effective strong dimension in algorithmic information and computational complexity. *SIAM Journal on Computing*. To appear.

[5] J. Berstel. Recent results in sturmian words. *Developments In Language Theory II*, pages 13–24, 1995.

[6] É. Borel. Sur les probabilités dénombrables et leurs applications arithmétiques. *Rend. Circ. Mat. Palermo*, 27:247–271, 1909.

[7] N. Chomsky and G. A. Miller. Finite state languages. *Information and Control*, 1:91–112, 1958.

[8] A. Cobham. Uniform tag sequences. *Mathematical Systems Theory*, 6:164–192, 1972.

[9] J. J. Dai, J. I. Lathrop, J. H. Lutz, and E. Mayordomo. Finite-state dimension. *Theoretical Computer Science*, 310(1–3):1–33, 2004.

[10] A. Ehrenfeucht, K.P. Lee, and G. Rozenberg. Subword complexities of various classes of deterministic developmental languages without interaction. *Theoretical Computer Science*, 1:59–75, 1975.

[11] K. Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons, second edition, 2003.

[12] F. Hausdorff. Dimension und äußeres Maß. *Mathematische Annalen*, 79:157–179, 1919.

[13] J. M. Hitchcock. Correspondence principles for effective dimensions. *Theory of Computing Systems*. To appear.

[14] J. M. Hitchcock. *Effective Fractal Dimension: Foundations and Applications*. PhD thesis, Iowa State University, 2003.

[15] J. M. Hitchcock. Fractal dimension and logarithmic loss unpredictability. *Theoretical Computer Science*, 304(1–3):431–441, 2003.

[16] J. M. Hitchcock and N. V. Vinodchandran. Dimension, entropy rates, and compression. In *Proceedings of the 19th IEEE Conference on Computational Complexity*, pages 174–183. IEEE Computer Society, 2004.

[17] W. Kuich. On the entropy of context-free languages. *Information and Control*, 16:173–200, 1970.

[18] M. Lothaire. *Combinatorics on Words*, volume 17 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Company, 1983.

[19] M. Lothaire. *Algebraic Combinatorics on Words*, volume 90 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 2002.

[20] J. H. Lutz. Dimension in complexity classes. *SIAM Journal on Computing*, 32(5):1236–1259, 2003.

[21] J. H. Lutz. The dimensions of individual strings and sequences. *Information and Computation*, 187(1):49–79, 2003.

[22] M. Rigo. Generalization of automatic sequences for numeration systems on a regular language. *Theoretical Computer Science*, 244:271–281, 2000.

[23] C. A. Rogers. *Hausdorff Measures*. Cambridge University Press, 1998. Originally published in 1970.

[24] C. P. Schnorr and H. Stimm. Endliche automaten und zufallsfolgen. *Acta Informatica*, 1:345–359, 1972.

[25] D. Sheinwald. On the Ziv-Lempel proof and related topics. *Proceedings of the IEEE*, 82(6):866–871, 1994.

[26] L. Staiger. Kolmogorov complexity and Hausdorff dimension. *Information and Computation*, 103:159–94, 1993.

[27] L. Staiger. A tight upper bound on Kolmogorov complexity and uniformly optimal prediction. *Theory of Computing Systems*, 31:215–29, 1998.

[28] D. Sullivan. Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups. *Acta Mathematica*, 153:259–277, 1984.

[29] C. Tricot. Two definitions of fractional dimension. *Mathematical Proceedings of the Cambridge Philosophical Society*, 91:57–74, 1982.

[30] J. Ziv and A. Lempel. Compression of individual sequences via variable rate coding. *IEEE Transactions on Information Theory*, 24:530–536, 1978.