Electronic Colloquium on Computational Complexity, Report No. 77 (2006)

# Lempel-Ziv Dimension for Lempel-Ziv Compression

Maria Lopez-Valdes*

**Abstract**

This paper describes the Lempel-Ziv dimension (Hausdorff like dimension inspired in the LZ78 parsing), its fundamental properties and relation with Hausdorff dimension. It is shown that in the case of individual infinite sequences, the Lempel-Ziv dimension matches with the asymptotical Lempel-Ziv compression ratio. This fact is used to describe results on Lempel-Ziv compression in terms of dimension of complexity classes and vice versa.

## 1 Introduction

Lutz [9] developed effective dimension (a Hausdorff like dimension) to quantitatively analyze the structure of complexity classes. Later, other authors have developed dimensions such as constructive or finite-state dimension and have found new connections with information theory. In particular, the motivation of this paper is the relation between dimension and compression. In this context, it was shown that polynomial-time and finite-state dimension can be characterized by the best compression ratio of polynomial-time and finite-state compressors respectively [8, 4]. As consequence, some results on dimension of complexity classes can be interpreted as compressibility results.

This paper focusses on the relation between dimension and the Lempel-Ziv compressor (LZ78) [12, 11]. This compression algorithm is probably the most widely studied of the universal compressors (compressors that do not depend on the distribution of the sequence source). Previous results have shown that polynomial-time and finite-state dimension are respectively a lower and upper bound of the asymptotical Lempel-Ziv compression ratio [8, 4]. However, by defining a dimension that matches the Lempel-Ziv compression ratio, results on Lempel-Ziv compression could be directly used to determine the dimension of some complexity classes. Alternatively, researchers in the compression domain have pointed out that it would be interesting to define a Hausdorff like dimension inspired in the LZ78 parsing, and to see in which cases this dimension would match with the Hausdorff

*Departamento de Informática e Ing. de Sistemas, María de Luna 1, Universidad de Zaragoza. 50018 Zaragoza, SPAIN. marlopez@unizar.es. This research was supported by Spanish Government MEC project TIN2005-08832-C03-02



ISSN 1433-8092

dimension. This is because it is expected that dimension would allow to deal with open problems related with Lempel-Ziv compression.

This paper describes the Lempel-Ziv dimension (Hausdorff like dimension inspired in the LZ78 parsing), its fundamental properties and relation with Hausdorff dimension. This dimension is naturally included in the existing hierarchy of dimensions defined until now. Furthermore, in the case of individual sequences, the Lempel-Ziv dimension matches with the asymptotical Lempel-Ziv compression ratio. This result is used to show some applications on dimension of complexity classes. Finally, results on dimension and ergodic theory are used to partially solve the open question raised by Lutz and others [6, 7] on the one-bit catastrophe in the LZ78 compressor.

This paper is distributed as follows. Section 2 outlines the preliminaries on the Lempel-Ziv compressor, measure, entropy and dimension. Section 3 describes the polynomial-time and finite-state dimension and some results related with compression. Section 4 develops the Lempel-Ziv dimension and fundamental properties. Section 5 describes results on Lempel-Ziv compression in terms of results on dimension of complexity classes and vice versa (in particular the one-bit catastrophe in the LZ78 compressor).

## 2 Preliminaries

Let a *string* be a finite and binary sequence $w \in \{0,1\}^*$. Let $|w|$ denote the length of a string and $\lambda$ the empty string. The *Cantor space* $\mathbf{C}$ is the set of all infinite binary sequences. Let $x[i \dots j]$ for $0 \leq i \leq j$ denote the $i$-th through the $j$-th bits of $x$, where $x \in \{0,1\}^* \cup \mathbf{C}$. Let $wx$ denote the concatenation of the string $w$ and the string or sequence $x$. Let $w \sqsubseteq x$ denote that $w$ is a *prefix* of $x$.

Let a *parsing* of a string $w \in \{0,1\}^*$ be a partition of $w$ into *phrases* $w_1, w_2, \dots, w_n$ such that $w_1 w_2 \dots w_n = w$. Let a *distinct parsing* of a string $w \in \{0,1\}^*$ be a parsing of $w$ such that no phrase, except possibly the last phrase, is the same as an earlier phrase. Let a *valid distinct parsing* of a string $w \in \{0,1\}^*$ be a distinct parsing of $w$ such that if $w_i$ is a phrase in the string $w$, then every prefix of $w_i$ appears before $w_i$ in the distinct parsing. Note that each string $w$ has an unique valid distinct parsing.

The LZ78 *compression algorithm* encodes a given string with its valid distinct parsing. This is done by replacing each phrase with a code word representing a pointer and a bit. The pointer indicates the longest proper prefix of the phrase and the bit is the last bit of the phrase. Together, they completely specify the phrase being encoded. The output of the compression algorithm on the string $w$ is denoted LZ($w$). (See [11, 12] for more details).

## 2.1 Probability measures on C and entropy

Let $C_w$ be the *cylinder* generated by a string $w \in \{0,1\}^*$, given by $C_w = \{S \in \mathbf{C} \mid w \sqsubseteq S\}$. Let $\mathcal{F}$ be the $\sigma$-algebra generated from the cylinder sets of $\mathbf{C}$. Let $\nu : \mathcal{F} \to [0,1]$ be a countable additive, nonnegative measure with total mass 1. Then, the triple $(\mathbf{C}, \mathcal{F}, \nu)$ is known as a *probability space* and $\nu$ is a *probability measure* on $\mathbf{C}$.

A measure $\nu$ is *stationary* if $\nu(T^{-1}X) = \nu(X)$ for all $X \in \mathcal{F}$, where $T$ is the left-shift on $\mathbf{C}$ (i.e. for $b \in \{0,1\}$ and $S \in \mathbf{C}$, $T(bS) := S$). A measure $\nu$ is *ergodic* if any $T$-invariant set $X$ (any set such that $T^{-1}X = X$) has $\nu(X) = 0$ or $\nu(X) = 1$.

On one hand, each probability measure $\nu$ is identified on $\mathbf{C}$ with a function $\mu : \{0,1\}^* \to [0,1]$ defined by $\mu(w) = \nu(C_w)$. This function $\mu$ verifies the following Kolmogorov's consistency conditions

$$
\begin{aligned}
&(i)\ \ \mu(\lambda) = 1.\\
&(ii)\ \mu(w0) + \mu(w1) = \mu(w) \ \text{ for all } w \in \{0,1\}^*.
\end{aligned}
$$

On the other hand, by Kolmogorov's Existence Theorem [1], for each $\mu : \{0,1\}^* \to [0,1]$ satisfying the consistence conditions, there exists a unique probability measure $\nu$ on $\mathbf{C}$ such that $\nu(C_w) = \mu(w)$. Then, for simplicity, a function $\mu$ verifying $(i)$ and $(ii)$ will be referred also as a probability measure on $\mathbf{C}$.

Let the *entropy* of an stationary measure $\mu$ on $\mathbf{C}$ be:

$$\mathcal{H}(\mu) = \lim_n \tfrac{\mathcal{H}_n(\mu)}{n}, \quad \text{where} \quad \mathcal{H}_n(\mu) = \textstyle\sum_{w \in \{0,1\}^n} \mu(w) \log \tfrac{1}{\mu(w)}.$$

**Theorem 2.1** *[3] (Entropy-rate Theorem) Let $\mu$ be a stationary ergodic probability measure on $\mathbf{C}$. Then, it exists a constant $h \geq 0$, such that*

$$\lim_n -\frac{1}{n} \log \mu(S[0 \ldots n-1]) = h$$

*almost surely. This constant $h$ is called the* entropy rate *of $\mu$.*

**Theorem 2.2** *[3] For a stationary measure $\mu$, the entropy $\mathcal{H}(\mu)$ is always defined. If $\mu$ is a stationary ergodic measure, then the entropy rate $h$ and the entropy $\mathcal{H}(\mu)$ are equal.*

### 2.2 Gales and Hausdorff dimension

**Definition.** Let $s \in [0, \infty)$.

1. An *$s$-supergale* is a function $d : \{0,1\}^* \to [0, \infty)$ satisfying

$$d(w) \geq 2^{-s}[d(w0) + d(w1)]$$

   for all $w \in \{0,1\}^*$.

2. An $s$-supergale $d$ *succeeds* on a sequence $S \in \mathbf{C}$ if

$$\limsup_n d(S[0 \ldots n]) = \infty.$$

3. The *success set* of $d$ is $S^\infty[d] = \{S \in \mathbf{C} \mid d \, \text{succeeds on} \, S\}$.

In 2003, Lutz [9] proved a characterization of classical Hausdorff dimension in terms of the $s$-supergales.

**Theorem 2.3** *[9] For every $X \subseteq \mathbf{C}$,*

$$\dim_H(X) = \inf\{s \in [0, \infty) \mid \exists \, s\text{-supergale} \, d \text{ s.t. } X \subseteq S^\infty[d]\}.$$

# 3 Polynomial-time dimension, Finite-State dimension and Lempel-Ziv compression

This section describes the polynomial-time and finite-state dimension and some results related with compression.

Based in the characterization of classical Hausdorff dimension, Lutz developed *resource-bounded dimension* [9] by introducing a resource-bound $\Delta$ and requiring the supergales to be $\Delta$-computable.

$$\dim_\Delta(X) = \inf\{s \in [0, \infty) \mid \exists \, \Delta\text{-computable } s\text{-supergale} \, d \text{ s.t. } X \subseteq S^\infty[d]\}.$$

This is the case of *polynomial-time dimension* that is defined by restricting attention to polynomial-time $s$-supergales (see [9] for further details).

**Definition.** The *polynomial-time dimension* of $X \subseteq \mathbf{C}$ is

$$\dim_{\mathrm{p}}(X) = \inf\{s \in [0, \infty) \mid \exists \, \mathrm{p}\text{-computable } s\text{-supergale} \, d \text{ s.t. } X \subseteq S^\infty[d]\}.$$

The polynomial-time dimension of a sequence $S \in \mathbf{C}$ is $\dim_{\mathrm{p}}(S) = \dim_{\mathrm{p}}(\{S\})$.

Also following this scheme, *finite-state dimension* is defined by restricting attention to $s$-supergales that are specified by finite-state devices (see [4] for a complete introduction).

**Definition.** The *finite-state dimension* of a set $X \subseteq \mathbf{C}$ is

$$\dim_{\mathrm{FS}}(X) = \inf\{s \in [0,\infty) \mid \exists \text{ finite-state } s\text{-supergale } d \text{ s.t. } X \subseteq S^{\infty}[d]\}.$$

The finite-state dimension of a sequence $S \in \mathbf{C}$ is $\dim_{\mathrm{FS}}(S) = \dim_{\mathrm{FS}}(\{S\})$.

The next proposition states that to add a string in the beginning of a sequence does not change its polynomial-time neither the finite-state dimension.

**Proposition 3.1** *For all $S \in \mathbf{C}$ and $w \in \{0,1\}^*$,*

1. $\dim_{\mathrm{p}}(wS) = \dim_{\mathrm{p}}(S)$.
2. $\dim_{\mathrm{FS}}(wS) = \dim_{\mathrm{FS}}(S)$.

The polynomial-time dimension is characterized as the best asymptotic compression-ratio attained by some special class of polynomial-time compressors. Lempel-Ziv compressor is in this class. Then,

**Proposition 3.2** *[8] For every $S \in C$,*

$$\dim_{\mathrm{p}}(S) \leq \liminf_{n} \frac{|\mathrm{LZ}(S[0\ldots n-1])|}{n}.$$

In [4] the authors obtained the following compressibility characterization of finite-state dimension of a sequence,

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

where $\rho_{\mathrm{FS}}(S)$ is the best compression ratio attainable for the infinite sequence $S$ by any information lossless finite-state compressor.

On the other side, it's well known [12, 11] that

$$\liminf_{n} \frac{|\mathrm{LZ}(S[0\ldots n-1])|}{n} \leq \rho_{\mathrm{FS}}(S).$$

Therefore, we can obtain the relationship between finite-state dimension and Lempel-Ziv compresion.

**Proposition 3.3** *For every $S \in \mathbf{C}$,*

$$\liminf_{n} \frac{|\mathrm{LZ}(S[0\ldots n-1])|}{n} \leq \dim_{\mathrm{FS}}(S).$$

Thus, the polynomial-time dimension and the finite-state dimension are respectively a lower and upper bound of the Lempel-Ziv compression-ratio.

## 4 Lempel-Ziv dimension

This section develops Lempel-Ziv dimension and its fundamental properties. Notice that the previous section is a good example of the close relationship between dimension and compression. The Lempel-Ziv dimension allows to see the Lempel-Ziv compression under the dimension point of view.

The definition of the Lempel-Ziv dimension is motivated by [10]. In this paper, the constructive dimension is defined by taking advantage of the existence of an universal constructive subprobability measure [13]. This allows us to center the definition in a single family of supergales $\{\tilde{d^s}\}_{s\geq 0}$,

$$\mathrm{cdim}(X) = \inf\{s \in [0,\infty) \mid X \subseteq S^{\infty}[\tilde{d^s}]\}.$$

The Lempel-Ziv algorithm (LZ78) is universal and asymptotically optimal for finite-state compressors [11, 12]. Therefore, as in the case of constructive dimension, we can define the Lempel-Ziv dimension in terms of a single family of supergales, $\{d^s_{\mathrm{LZ}}\}_{s\geq 0}$.

**Definition.** For each $s \in [0,\infty)$, let the *Lempel-Ziv s-supergale* $d^s_{\mathrm{LZ}}$ be:

$$\begin{aligned} d^s_{\mathrm{LZ}}(\lambda) &= 1 \\ d^s_{\mathrm{LZ}}(w) &= \begin{cases} \frac{2^{s|w|}}{n!}\frac{\#\{i\in\{1\ldots n\} \mid u\sqsubseteq w_i\}}{n} & \text{if} \quad w = w_1w_2\ldots w_n u \\ \frac{2^{s|w|}}{n!} & \text{if} \quad w = w_1w_2\ldots w_n \end{cases} \end{aligned}$$

where $w_1,\ldots w_n$ are the distinct phrases in the valid distinct parsing of $w$ and $u = w_i$ for some $i \in \{1\ldots n\}$.

**Observation 4.1** *For every polynomial-time computable real $s$, the Lempel-Ziv $s$-supergale is polynomial-time computable.*

**Proposition 4.2** *The Lempel-Ziv $s$-supergale is optimal for the class of finite-state $s$-supergales. That is, there exists $\alpha > 0$ such that for all $s \in [0,\infty)$ and all $d$ finite-state $s$-supergale,*

$$d^s_{\mathrm{LZ}}(w) \geq \alpha d(w),$$

*for every $w \in \{0,1\}^*$ (long enough).*

**Observation 4.3** *For all $s,t \in [0,\infty)$ and $w \in \{0,1\}^*$,*

$$d^s_{\mathrm{LZ}}(w)2^{-s|w|} = d^t_{\mathrm{LZ}}(w)2^{-t|w|}.$$

**Definition.** The *Lempel-Ziv dimension* of $X \subseteq \mathbf{C}$ is,

$$\dim_{\mathrm{LZ}}(X) = \inf\{s \in [0, \infty) \mid X \subseteq S^{\infty}[d^s_{\mathrm{LZ}}]\}.$$

The *Lempel-Ziv dimension* of a sequence $S \in \mathbf{C}$ is $\dim_{\mathrm{LZ}}(S) = \dim_{\mathrm{LZ}}(\{S\})$.

**Observation 4.4** *For all* $X \subseteq Y \subseteq \mathbf{C}$*,* $\dim_{\mathrm{LZ}}(X) \leq \dim_{\mathrm{LZ}}(Y)$*.*

The following theorem states that the Lempel-Ziv dimension of any set $X \subseteq \mathbf{C}$ is completely determined by the dimension of the individual sequences in the set.

**Theorem 4.5** *For all* $X \subseteq \mathbf{C}$*,*

$$\dim_{\mathrm{LZ}}(X) = \sup_{S \in X} \dim_{\mathrm{LZ}}(S).$$

This theorem implies one important property of dimension, its *countable stability.*

**Corollary 4.6**

1. *For all sets* $X, Y \subseteq \mathbf{C}$*,*

$$\dim_{\mathrm{LZ}}(X \cup Y) = \max\{\dim_{\mathrm{LZ}}(X), \dim_{\mathrm{LZ}}(Y)\}.$$

2. *Let* $X_1, X_2 \ldots \subseteq \mathbf{C}$*,*

$$\dim_{\mathrm{LZ}}(\bigcup_{i=1}^{\infty} X_i) = \sup_{i \in \mathbb{N}} \dim_{\mathrm{LZ}}(X_i).$$

The main theorem of this section gives an exact characterization of the Lempel-Ziv dimension of an infinite sequence in terms of the asymptotical compression ratio attained by the Lempel-Ziv algorithm.

**Theorem 4.7** *Let* $S \in \mathbf{C}$*,*

$$\dim_{\mathrm{LZ}}(S) = \liminf_{n} \frac{|\mathrm{LZ}(S[0 \ldots n-1])|}{n}.$$

The following result is a consequence of Observation 4.1 and Proposition 4.2. By Theorem 4.7, the second part is a reformulation of Propositions 3.2 and 3.3 in terms of dimension.

**Theorem 4.8** *Let $X \subseteq \mathbf{C}$,*

$$\dim_{\mathrm{p}}(X) \leq \dim_{\mathrm{LZ}}(X) \leq \dim_{\mathrm{FS}}(X).$$

*In particular, for all $S \in \mathbf{C}$,*

$$\dim_{\mathrm{p}}(S) \leq \dim_{\mathrm{LZ}}(S) \leq \dim_{\mathrm{FS}}(S).$$

By Proposition 3.1 and Theorem 4.8 we have the following relationship between the Lempel-Ziv dimension of $S$ and $wS$.

**Theorem 4.9** *Let $S \in \mathbf{C}$ and $w \in \{0,1\}^*$, then*

$$|\dim_{\mathrm{LZ}}(S) - \dim_{\mathrm{LZ}}(wS)| \leq \dim_{\mathrm{FS}}(S) - \dim_{\mathrm{p}}(S).$$

*In particular, if $\dim_{\mathrm{FS}}(S) = \dim_{\mathrm{p}}(S)$, then for all $w \in \{0,1\}^*$,*

$$\dim_{\mathrm{LZ}}(S) = \dim_{\mathrm{LZ}}(wS).$$

A consequence of this Theorem is that for sequences such that the polynomial-time dimension and finite-state dimension are equal, the one-bit catastrophe (defined in the next section) is not verified.

**Definition.** Let the class LZBIT be the set of all sequences $S$ such that for all $w \in \{0,1\}^*$, $\dim_{\mathrm{LZ}}(S) = \dim_{\mathrm{LZ}}(wS)$.

We use Lempel-Ziv dimension to endow LZBIT with internal dimension structure.

**Definition.** For $X \subseteq \mathbf{C}$, the *dimension of $X$* in LZBIT is

$$\dim(X \mid \mathrm{LZBIT}) = \dim_{\mathrm{LZ}}(X \cap \mathrm{LZBIT}).$$

# 5 Applications

In this section we partially solve the open question raised by Jack Lutz and other authors [6, 7] on the one-bit catastrophe (without the need of studying the asymptotic valid distinct parsing). We also give some results about the Lempel-Ziv compressibility of sequences from results in polynomial-time and finite-state dimension and vice versa.

The one-bit catastrophe conjecture says that the compression ratio of an infinite sequence can change substantially when we add an initial bit on the sequence. In terms of dimension the conjecture reads as: $S \in \mathbf{C}$ verifies the one-bit catastrophe iff $\dim_{\mathrm{LZ}}(S) \neq \dim_{\mathrm{LZ}}(bS)$ for some $b \in \{0,1\}$. Let us illustrate this conjecture with the following example.

**Example 5.1** *Let* $S = 101100111000\dots 1^n 0^n \dots \in \mathbf{C}$ *that is clearly highly compressed by Lempel-Ziv algorithm. Let* $t(w)$ *be the number of phrases in the valid distinct parsing of* $w \in \{0,1\}^*$. *Let* $w_1 = S[0\dots 29]$, $w_2 = S[0\dots 109]$ *and* $w_3 = S[0\dots 239]$. *Then,*

$$\begin{array}{llll} t(w_1) = 10 \Rightarrow |\mathrm{LZ}(w_1)| = 35 & & t(1w_1) = 13 \Rightarrow |\mathrm{LZ}(1w_1)| = 53 \\ t(w_2) = 20 \Rightarrow |\mathrm{LZ}(w_2)| = 101 & & t(1w_2) = 29 \Rightarrow |\mathrm{LZ}(1w_2)| = 146 \\ t(w_3) = 30 \Rightarrow |\mathrm{LZ}(w_3)| = 151 & & t(1w_3) = 45 \Rightarrow |\mathrm{LZ}(1w_3)| = 271 \end{array}$$

*In the case of* $S$ *it seems that the longer the prefixes are, the better they are compressed. However, in the case of the prefixes of* $1S$, *it seems that LZ78 does not compress anything.*

*With results derived in this section, we will show that both sequences are asymptotically highly compressible and do not verify the one-bit catastrophe.*

## 5.1 Stochastic sequences and ergodic measures

In this subsection we present classes of stochastic sequences that do not verify the one-bit catastrophe with probability 1 and determine its Hausdorff, polynomial-time and Lempel-Ziv dimension.

**Definition.** Let $S \in \mathbf{C}$ and $m \in \mathbb{N}$. Let the *relative probability* for each $w \in \{0,1\}^m$ and each $n \geq m$ be

$$p_m(w|S[0\dots n-1]) = \frac{\#\{0 \leq i \leq n+m \mid S[i\dots i+m-1] = w\}}{n-m+1}.$$

The *limiting relative probability* is defined as

$$p_m(w|S) = \lim_n p_m(w|S[0\dots n-1]),$$

provided the limit exists.

**Definition.** A sequence $S \in \mathbf{C}$ is *stochastic* if $p_m(w|S)$ exists for any $m \in \mathbb{N}$ and any $w \in \{0,1\}^m$. Let $\mathcal{S} \subseteq \mathbf{C}$ denote the set of all stochastic sequences.

Every stochastic sequence $S \in \mathcal{S}$ induces a unique stationary measure on $\mathbf{C}$, $\mu_S : \{0,1\}^* \to [0,1]$ such that $\mu_S(w) = p_{|w|}(w|S)$. (See [1]).

**Definition.** Let $\mu$ be a stationary measure. The set of *frequency typical sequences* of $\mu$ is:

$$\mathcal{T}(\mu) = \{S \in \mathcal{S} \mid \mu_S = \mu\}.$$

**Observation 5.2** *Let $S \in \mathcal{S}$ and $w \in \{0,1\}^*$, then $wS \in \mathcal{S}$ and $wS \in \mathcal{T}(\mu_S)$.*

The next results will be useful to determine the Lempel-Ziv dimension of frequency typical sequences of a stationary ergodic measure.

**Theorem 5.3** *(LZ78 Universality Theorem) Let $\mu$ be a stationary ergodic measure on $\mathbf{C}$ with entropy rate $h$ and $S \in \mathcal{T}(\mu)$ then,*

$$\lim_n \frac{|\mathrm{LZ}(S[0 \ldots n-1])|}{n} = h \qquad \textit{almost surely.}$$

From [2] we have the following result.

**Proposition 5.4** *Let $\mu$ be a stationary ergodic measure with entropy rate $h$. For all $S \in \mathcal{T}(\mu)$,*

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

**Theorem 5.5** *[5] Let $\mu$ be a stationary ergodic measure with entropy rate $h$. The Hausdorff dimension of the set of frequency typical sequences of $\mu$ is $h$. That is,*

$$\dim_H(\mathcal{T}(\mu)) = h.$$

**Corollary 5.6** *Let $\mu$ be a stationary ergodic measure with entropy rate $h$, then*

1. $\dim_{\mathrm{p}}(\mathcal{T}(\mu)) = \dim_{\mathrm{LZ}}(\mathcal{T}(\mu)) = h$.
2. $\dim(\mathcal{T}(\mu) \,|\, \mathrm{LZBIT})) = h$.
3. *If $S \in \mathcal{T}(\mu)$, then $\dim_{\mathrm{LZ}}(S) = h$ almost surely.*
4. *If $S \in \mathcal{T}(\mu)$, then $S \in \mathrm{LZBIT}$ almost surely.*

This corollary states that: $(i)$ almost all frequency typical sequences of a stationary ergodic measure have Lempel-Ziv compression ratio equal to the entropy rate; $(ii)$ the compression ratio of all of them is always less than $h$; and $(iii)$ almost all of them do not verify the one-bit catastrophe.

We use next the last Corollary in a particular class of measures, which includes the uniform measure.

**Definition.** Let $\alpha \in [0,1]$. The *$\alpha$-coin-toss probability measure* on $\mathbf{C}$ is

$$\mu^\alpha(w) = (1-\alpha)^{\#(0,w)} \alpha^{\#(1,w)},$$

where $\#(b, w)$ is the number of times that the bit $b$ appears in the string $w$. In other words, $\mu^\alpha(w)$ is the probability that $S \in C_w$ when $S \in \mathbf{C}$ is chosen according to a random experiment in which the $i^{\text{th}}$ bit of $S$ is decided by tossing a 0/1-valued coin whose probability of 1 is $\alpha$.

**Proposition 5.7** *Let $\alpha \in [0,1]$ and let $\mathcal{H}$ be the* binary entropy function $\mathcal{H} : [0,1] \to [0,1]$ *defined by*

$$\mathcal{H}(\alpha) = \alpha \log \frac{1}{\alpha} + (1-\alpha) \log \frac{1}{1-\alpha}.$$

*Then, $\mu^\alpha$ is a stationary ergodic measure with entropy rate $\mathcal{H}(\alpha)$.*

**Definition.** A sequence $S \in \mathbf{C}$ is normal ($S \in$ NORMAL) if $S \in \mathcal{T}(\mu^{\frac{1}{2}})$. That is, $S$ is normal if every string $w \in \{0,1\}^*$ has asymptotic frequency $2^{-|w|}$ in $S$.

**Theorem 5.8** *Let $\alpha \in [0,1]$.*

1. $\dim_{\mathrm{H}}(\mathcal{T}(\mu^\alpha)) = \dim_{\mathrm{p}}(\mathcal{T}(\mu^\alpha)) = \dim_{\mathrm{LZ}}(\mathcal{T}(\mu^\alpha)) = \mathcal{H}(\alpha)$.
2. $\dim(\mathcal{T}(\mu^\alpha) \mid \mathrm{LZBIT}) = \mathcal{H}(\alpha)$.
3. *For $S \in \mathcal{T}(\mu)$,* $\dim_{\mathrm{LZ}}(S) = \mathcal{H}(\alpha)$ *almost surely.*
4. *For $S \in \mathcal{T}(\mu)$, $S \in$ LZBIT almost surely.*

*In particular,*

1. $\dim_{\mathrm{H}}(\mathrm{NORMAL}) = \dim_{\mathrm{p}}(\mathrm{NORMAL}) = \dim_{\mathrm{LZ}}(\mathrm{NORMAL}) = 1$
2. $\dim(\mathrm{NORMAL} \mid \mathrm{LZBIT}) = 1$.
3. *For $S \in$ NORMAL,* $\dim_{\mathrm{LZ}}(S) = 1$ *almost surely.*
4. *For $S \in$ NORMAL, $S \in$LZBIT almost surely.*

In [7] it is proved that sequences with Lempel-Ziv dimension 1 (that is, sequences that Lempel-Ziv algorithm does not compress) are in NORMAL. However, it is also showed that the converse does not hold. Notice that here we show that almost all normal sequences have $\dim_{\mathrm{LZ}}(S) = 1$ and do not verify the one-bit catastrophe.

Since the binary entropy function is surjective on [0,1], then we have that there exist sequences with Lempel-Ziv compression ratio $\eta$ (for any $\eta \in [0,1]$) that don't verify the one-bit catastrophe:

**Corollary 5.9** *For all $\eta \in [0,1]$ there exist sequences $S \in$ LZBIT with* $\dim_{\mathrm{LZ}}(S) = \eta$.

### 5.2 Highly LZ compressible sequences.

**Definition.** An infinite sequence $S \in \mathbf{C}$ is highly LZ compressible if

$$\liminf_n \frac{|\mathrm{LZ}(S[0\ldots n-1])|}{n} = 0$$

That is, in terms of Lempel-Ziv dimension, $\dim_{\mathrm{LZ}}(S) = 0$.

**Definition.** Let $S \in \mathbf{C}$ and $m \in \mathbb{Z}^+$,

1. The *factor set* $F_m(S)$ is the set of all finite strings of length $m$ that appear in $S$.

2. The *factor complexity function*, $p_S : \mathbb{N} \to \mathbb{N}$, counts the number of factors of each $m$, that is $p_S(m) = |F_m(S)|$.

In [2] it is proved that finite-state dimension of sequences with $\mathrm{p}_S(m) = 2^{o(m)}$ are equal to zero. Then we have the following theorem.

**Theorem 5.10** *Every $S \in \mathbf{C}$ with $\mathrm{p}_S(m) = 2^{o(m)}$ is highly LZ compressible and $S \in$ LZBIT.*

In particular, by using this result on example 5.1, sequence $S$ is highly compressible and $S \in$ LZBIT since $S$ satisfies $\mathrm{p}_S(m) = m(m+1)$.

Other applications of this Theorem are,

**Corollary 5.11**

1. *If $S$ is the binary expansion of a rational number, $S$ is highly LZ compressible and $S \in$ LZBIT.*

2. *Sturmian sequences, Morphic sequences, Automatic sequences are highly LZ compressible and are in LZBIT (See [2]).*

3. *Every $S \in$ REG is highly LZ compressible and $S \in$ LZBIT.*

## References


[1] P. Billingsley. *Probability and Measure.* John Wiley & Sons, Inc., New York, N.Y., 1979.

[2] J. M. Hitchcock C. Bourke and N. V.Vinodchandran. Entropy rates and finite-state dimension. *Theoretical Computer Science*, 349, 2004.

[3] T.M. Cover and J.A. Thomas. *Elements of Information Theory.* John Wiley & Sons, Inc., New York, N.Y., 1991.

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

[5] K. Hojo, B. Ryabko, and J. Suzuki. Performance of data compression in terms of hausdorff dimension. *TIEICE: IEICE Transactions on Communications/Electronics/ Information and Systems*, 310:1761–1764, 2001.

[6] P.C. Shields L.A. Pierce II. Sequences incompressible by SLZ (LZW) yet fully compressible by ULZ. *Numbers, Information and Complexity, Kluwer Academic Publishers*, pages 385–390, 2000.

[7] J. I. Lathrop and M. J. Strauss. A universal upper bound on the performance of the Lempel-Ziv algorithm on maliciously-constructed data. In B. Carpentieri, editor, *Compression and Complexity of Sequences '97*, pages 123–135. IEEE Computer Society Press, 1998.

[8] M. López-Valdés and E. Mayordomo. Dimension is compression. In *Proceedings of the 30th International Symposium on Mathematical Foundations of Computer Science*, pages 676–685. Springer-Verlag, 2005.

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

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

[11] D. Scheinwald. On the lempel-ziv proof and related topics. In *Proceedings of the IEEE*, volume 82, pages 866–871, 1994.

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

[13] A. K. Zvonkin and L. A. Levin. The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematical Surveys*, 25:83–124, 1970.

## Appendix

**Proof of Proposition 4.2.** We will need the following results and definitions in our proof.

*Definition.* A *finite-state compressor (FSC)* is a 4-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*.
- $q_0$ is the *initial state*.

For $q \in Q$ and $w \in \{0,1\}^*$, we define the *output* from state $q$ on input $w$ to be the string $\nu(q, w)$ defined by recursion

$$\begin{aligned} \nu(q, \lambda) &= \lambda \\ \nu(q, wb) &= \nu(q, w)\nu(\delta^*(q, w), b) \end{aligned}$$

for all $w \in \{0,1\}^*$ and $b \in \{0,1\}$. Here $\delta^* : Q \times \{0,1\}^* \to Q$ is defined by the recursion

$$\begin{aligned} \delta^*(q, \lambda) &= q \\ \delta^*(q, wb) &= \delta(\delta^*(q, w), b) \end{aligned}$$

for all $q \in Q$, $w \in \{0,1\}^*$ and $b \in \{0,1\}$; we write $\delta(w)$ for $\delta^*(q_0, w)$. We then define the *output* of $C$ on input $w \in \{0,1\}^*$ to be the string $C(w) = \nu(q_0, w)$.

*Definition.* An FSC $C = (Q, \delta, \nu, q_0)$ is *information-lossless (IL)* if the function

$$\begin{aligned} \{0,1\}^* &\to \{0,1\}^* \times Q \\ w &\mapsto (C(w), \delta(w)) \end{aligned}$$

is one-to-one. An *information-lossless finite-state compressor (ILFSC)* is an FSC that is IL.

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

- $Q$ is 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 $q \in Q$.
- $q_0$ is the *initial state*.

We extend $\delta$ to the transition function $\delta : \{0,1\}^* \to Q$ that is defined as before.

*Definition.* Let $G = (Q, \delta, \beta, q_0)$ be a finite-state gambler, the *martingale of G* is the function $d_G : \{0,1\}^* \to [0,\infty)$ defined by the recursion

$$\begin{aligned} d_G(\lambda) &= 1 \\ d_G(wb) &= 2d_G(w)[(1-b)(1-\beta(\delta(w))) + b\beta(\delta(w))] \end{aligned}$$

for all $w \in \{0,1\}^*$ and $b \in \{0,1\}$.

*Definition.* For $s \in [0,\infty)$ a *finite-state s-gale d* is an $s$-gale for which there exists an FSG $G$ such that $d = 2^{(s-1)|w|} d_G(w)$.

*Lemma.* Let $G = (Q, \delta, \beta, q_0)$ be a nonvanishing 1-account finite-state gambler (FSG). Then, for every $w \in \{0,1\}^*$ with $w_0 \dots w_{n-1}$ its valid distinct parsing, there exists $C = C(w,G) = (Q', \delta', \nu', q_0')$ ILFSC such that

$$|C(w)| \leq n + |w| - \log d_G(w)$$

*Proof.* Let $k_i = |w_i|$ and $k = \max\{k_i \mid 1 \leq i \leq n\}$. Let $q_i = \delta(w_0 \dots w_i)$ and $G_i = (Q, \delta, \beta, q_i)$. Define $p_i : \{0,1\}^{k_i} \to [0,1]$ by

$$p_i(u) = 2^{-k_i} d_{G_i}(u).$$

Since $d_{G_i}$ is a martingale with $d_{G_i}(\lambda) = 1$, each of the functions $p_i$ is a positive probability measure on $\{0,1\}^{k_i}$.

For each $q_i \in Q$ let $\mathfrak{S}_\mathfrak{i} : \{0,1\}^{k_i} \to \{0,1\}^*$ be the Shannon code (see [3]) given by the probability measure $p_i$. Then,

$$\mathfrak{S}_\mathfrak{i}(u) = \lceil \log \frac{1}{p_i(u)} \rceil.$$

For all $u \in \{0,1\}^{k_i}$ and $i \in \{0, \dots n-1\}$. We define the FSC $C = (Q', \delta', \nu', q_0')$ whose components are as follows.

(i) $Q' = Q \times \{0, 1, \ldots n\} \times \{0, 1\}^{<k}$.

(ii) For all $(q, m, u) \in Q'$ and $b \in \{0, 1\}$,

$$\delta'((q, m, u), b) = \begin{cases} (q, m, ub) & \text{if } |u| < k_{m+1} - 1 \\ (\delta(q, ub), m+1, \lambda) & \text{if } |u| = k_{m+1} - 1 \end{cases}$$

(iii) For all $(q, m, u) \in Q'$ and $b \in \{0, 1\}$,

$$\nu'((q, m, u), b) = \begin{cases} \lambda & \text{if } |u| < k_{m+1} - 1 \\ \mathfrak{S}_{\mathfrak{i}}(ub) & \text{if } |u| = k_{m+1} - 1 \end{cases}$$

(iv) $q_0' = (q_0, 0, \lambda)$.

Since $\mathfrak{S}_{\mathfrak{i}}$ is a uniquely decodable code, its easy to see that the FSC $C$ is IL. Now, $C(w) = \mathfrak{S}_{\mathfrak{o}}(w_0) \ldots \mathfrak{S}_{\mathfrak{n-1}}(w_{n-1})$ and then

$$\begin{aligned} |C(w)| &= \sum_{i=0}^{n-1} |\mathfrak{S}_{\mathfrak{i}}(w_i)| = \sum_{i=0}^{n-1} \left\lceil \log \frac{1}{p_i(w_i)} \right\rceil \\ &\leq \sum_{i=0}^{n-1} \left( 1 + \log \frac{2^{|w_i|}}{d_{G_i}} \right) = n + |w| - \log d_G(w). \end{aligned}$$

□

Let $d$ be a finite-state $s$-gale, by definition there exists a nonvanishing 1-account finite-state gambler (FSG) $G$ such that $d(w) = 2^{(s-1)|w|} d_G(w)$. By the previous lemma, for each $w$ there exists $C = C(w, G)$ ILFSC such that

$$d(w) \leq 2^{s|w| + n - |C(w)|}$$

So, it's suffices to prove that for all $w \in \{0, 1\}^*$ there exists $\alpha > 0$ (not depending on $w$) such that for all $C = C(w, G)$,

$$\alpha 2^{s|w| + n - |C(w)|} \leq d_{\mathrm{LZ}}(w)$$

Let $w$ such that $w_1, w_2 \ldots w_n$ is the valid distinct parsing of $w$. Then $d_{\mathrm{LZ}}(w) \geq \frac{2^{s|w|}}{n!} \cdot \frac{1}{n}$. By the following formula

$$\log n! = n \log n - n + O(\log n),$$

we only need to show that

$$\alpha 2^{n - |C(w)|} \leq 2^{-n \log n - O(\log n)}$$

It is known [12] that

$$|C(w)| \geq (n+\xi^2)\log\frac{n+\xi^2}{4\xi^2} + 2\xi^2$$

where $|Q'| = \xi$. And there exists $\alpha > 0$ such that for $n \in \mathbb{N}$ long enough we have that

$$\alpha 2^{-(n+\xi^2)\log\frac{n+\xi^2}{4\xi^2} - 2\xi^2} \leq 2^{-n\log n - O(\log n)}$$

Then, for $w$ long enough, we have that $d_{\mathrm{LZ}}$ is optimal in the class of finite-state $s$-gales.

**Proof of Theorem 4.5.** Let $X \subseteq \mathbf{C}$ and let $s' = \sup_{S\in X}\dim_{\mathrm{LZ}}(S)$. It is clear by Observation 4.4 that $\dim_{\mathrm{LZ}}(X) \geq s'$.

To see that $\dim_{\mathrm{LZ}}(X) \leq s'$, let $s$ be a real number such that $s > s'$, it sufficies to show that $\dim_{\mathrm{LZ}}(X) \leq s$.

Since $s > s'$, for each $S \in X$, $d^s_{\mathrm{LZ}}$ succeeds on $S$ and this proves that $\dim_{\mathrm{LZ}}(X) \leq s$.

**Proof of Theorem 4.7.** We first prove the following lemma.

*Lemma.* Let $S \in \mathbf{C}$,

$$\dim_{\mathrm{LZ}}(S) = \inf\{s \in [0,\infty) \mid \limsup_n d^s_{\mathrm{LZ}}(S[0\ldots n-1]) > 1\}.$$

*Proof.* By definition of $\dim_{\mathrm{LZ}}$ we have

$$\dim_{\mathrm{LZ}}(S) \geq \inf\{s \in [0,\infty) \mid \limsup_n d^s_{\mathrm{LZ}}(S[0\ldots n-1]) > 1\}.$$

To see the other inequality, let

$$s' = \inf\{s \in [0,\infty) \mid \limsup_n d^s_{\mathrm{LZ}}(S[0\ldots n-1]) > 1\}$$

and let $t' > t > s'$. Then by Observation 4.3,

$$\limsup_n d^{t'}_{\mathrm{LZ}}(S[0\ldots n-1]) = \limsup_n 2^{(t'-t)n} d^{t}_{\mathrm{LZ}}(S[0\ldots n-1]) = \infty$$

□

To prove that $\dim_{\mathrm{LZ}}(S) \geq \liminf_n \frac{|\mathrm{LZ}(S[0\ldots n-1])|}{n}$, let $s \geq \dim_{\mathrm{LZ}}(S)$. By the previous lemma,

$$\limsup_n d^s_{\mathrm{LZ}}(S[0\ldots n-1]) > 1$$

and there exist infinitely many $n \in \mathbb{N}$ such that $d^s_{\mathrm{LZ}}(S[0\ldots n-1]) > 1$. That is,

$$\exists^\infty n \qquad \frac{2^{sn}}{z(S[0\ldots n-1])!} \cdot \frac{\rho(S,n)}{z(S[0\ldots n-1])} > 1 \tag{1}$$

where $z(S[0\ldots n-1])$ denotes the number of different phrases in the valid distinct parsing of $S[0\ldots n-1]$ and $\rho(S,n)$ is the number of phrases in $S[0\ldots n-1]$ such that are extensions of the last phrase.

From (1), there exist infinitely many $n \in \mathbb{N}$ such that

$$\begin{aligned} s &> \frac{\log(z(S[0\ldots n-1])(z(S[0\ldots n-1]))!) - \log(\rho(S,n))}{n} \\ &\geq \frac{\sum_{k=1}^{z(S[0\ldots n-1])} \log k}{n} - \frac{\log(\rho(S,n))}{n} \end{aligned}$$

Since $\rho(S,n) \leq n$, taking $n$'s large enough, there exist infinitely many $n \in \mathbb{N}$ such that

$$s > \frac{\sum_{k=1}^{z(S[0\ldots n-1])} \log k}{n}. \tag{2}$$

It is known [3] the following bound on the number of phrases $t(n)$ in a distinct parsing of a binary sequence of length $n$:

$$t(n) \leq \frac{n}{(1-\epsilon_n)\log n} \tag{3}$$

where $\epsilon_n \to 0$ as $n \to \infty$.

Then, by (2) and (3)

$$\begin{aligned} s > \quad & \liminf_n \frac{\sum_{k=1}^{z(S[0\ldots n-1])} \log k}{n} = \\ & \liminf_n \frac{\sum_{k=1}^{z(S[0\ldots n-1])} \lceil \log k \rceil + z(S[0\ldots n-1])}{n} = \\ & \liminf_n \frac{|\mathrm{LZ}(S[0\ldots n-1])|}{n} \end{aligned}$$

which prove the inequality

$$\dim_{\mathrm{LZ}}(S) \geq \liminf_n \frac{|\mathrm{LZ}(S[0\ldots n-1])|}{n}$$

To see the other inequality, let $s > \liminf_n \frac{|\mathrm{LZ}(S[0\ldots n-1])|}{n}$. In the same manner that before, we can see that there exist infinitely many $n \in \mathbb{N}$ such that

$$s > \frac{\log(z(S[0\ldots n-1])!) - \log(\rho(S,n))}{n}$$

From this,

$$\begin{aligned}
&\limsup_n d^s_{\mathrm{LZ}}(S[0\ldots n-1]) \geq \\
&\liminf_n d^s_{\mathrm{LZ}}(S[0\ldots n-1]) \geq \\
&\liminf_n \frac{2^{sn}}{z(S[0\ldots n-1])!}\rho(S,n) > 1
\end{aligned}$$

and $s \leq \dim_{\mathrm{LZ}}(S)$, which proves the theorem.

**Proof of Theorem 4.9.** Let $S \in \mathbf{C}$ and $w \in \{0,1\}^*$.
If $\dim_{\mathrm{LZ}}(S) = \dim_{\mathrm{LZ}}(wS)$ the proof is obvious.
If $\dim_{\mathrm{LZ}}(S) > \dim_{\mathrm{LZ}}(wS)$ then by Theorem 4.8 and Proposition 3.1

$$\begin{aligned}
\dim_{\mathrm{LZ}}(S) - \dim_{\mathrm{LZ}}(wS) &\leq \dim_{\mathrm{FS}}(S) - \dim_{\mathrm{p}}(wS) \\
&= \dim_{\mathrm{FS}}(S) - \dim_{\mathrm{p}}(S).
\end{aligned}$$

If $\dim_{\mathrm{LZ}}(wS) > \dim_{\mathrm{LZ}}(S)$ similarly,

$$\begin{aligned}
\dim_{\mathrm{LZ}}(S) - \dim_{\mathrm{LZ}}(wS) &\leq \dim_{\mathrm{FS}}(wS) - \dim_{\mathrm{p}}(S) \\
&= \dim_{\mathrm{FS}}(S) - \dim_{\mathrm{p}}(S).
\end{aligned}$$

**Proof of Theorem 5.3.** Let $\mu$ be a stationary ergodic measure on $\mathbf{C}$ with entropy rate $h$ and $S \in \mathcal{T}(\mu)$. In [12] it is proved that

$$\limsup_n \frac{|LZ(S[0\ldots n-1])|}{n} \leq h$$

almost surely. To see the other inequality, let

$$A_n = \{w \in \{0,1\}^n \mid |\mathrm{LZ}(w)| \leq -\log\mu(w) - \log n\}.$$

The measure of $A_n$ is

$$\mu(A_n) = \sum_{w \in A_n} \mu(w) \leq 2^{-\log n} \sum_{w \in A_n} 2^{-|\mathrm{LZ}(w)|} \leq \frac{1}{n}$$

where the last inequality it follows from Kraft's inequality [3]. Then, $\mu(A_n)$ tends to zero with $n$ and

$$|LZ(w)| + \log n \geq -\log\mu(w), \tag{4}$$

almost surely.
Let $S \in \mathcal{T}(\mu)$, then by (4)

$$\liminf_n \frac{|LZ(S[0\ldots n-1])|}{n} \geq \liminf \frac{-\log \mu(S[0\ldots n-1])}{n} = h$$

almost surely.

**Proof of Proposition 5.4.** We will need the following result by [2].

*Theorem.* For every $S \in \mathbf{C}$, $\dim_{\mathrm{FS}}(S) = H(S)$, where $H(S)$ is the block-entropy rate of $S$.

From the definition of block-entropy rate of $S$ (see [2]), if $S$ is an stochastic sequence then $H(S) = H(\mu_S)$.

Let $\mu$ be a stationary ergodic measure with entropy rate $h$ and let $S \in \mathcal{T}(\mu)$. Then,

$$H(S) = H(\mu_S) = H(\mu) = h.$$

**Proof of Proposition 5.5.** It is well known that

$$\dim_{\mathrm{H}}(X) \leq \dim_{\mathrm{p}}(X) \leq \dim_{\mathrm{FS}}(X) \tag{5}$$

Let $\mu$ be an ergodic measure with entropy rate $h$. By Proposition 5.4 and (5), $\dim_{\mathrm{H}}(X) \leq h$.

We next prove that the other inequality holds. Let $\epsilon > 0$ and let $A_n$ be the typical set with respect $\mu$,

$$A_n = \{w \in \{0,1\}^n \mid 2^{-n(h+\epsilon)} \leq \mu(w) \leq 2^{-n(h-\epsilon)}\} \tag{6}$$

The number of elements in $A_n$ is at least $(1-\epsilon)2^{n(h-\epsilon)}$. This follows from the Entropy-rate Theorem, for any $\delta > 0$, there exists an $n_0$ such that for all $n \geq n_0$, we have

$$\mu(A_n) = \mu(\{w \in \{0,1\}^n \mid |-\frac{1}{n}\log \mu(w) - h| \leq \epsilon\}) > 1 - \delta. \tag{7}$$

Setting $\delta = \epsilon$ we obtain that, for $n$ sufficiently large, $\mu(A_n) > 1-\epsilon$, so that

$$\begin{aligned} 1-\epsilon &< \mu(A_n) \leq \sum_{w \in A_n} \mu(w) \\ &\leq \sum_{w \in A_n} 2^{-n(h-\epsilon)} = 2^{-n(h-\epsilon)}|A_n| \end{aligned}$$

Hence

$$|A_n| \geq (1-\epsilon)2^{n(h-\epsilon)} \tag{8}$$

Let $d$ be an $s$-supergale that succeeds on $\mathcal{T}(\mu)$. Without loss of generality we can assume that $d(\lambda) = 1$. Let

$$B_n = \{w \in \{0,1\}^n \mid d(w) > 1\}$$

Let $S \in \mathcal{T}(\mu)$. By the Entropy-rate Theorem, for $n$ sufficiently large, $S[0 \ldots n-1] \in A_n$. Since $S \in S^\infty[d]$, it follows that there exist infinitely many $n \in \mathbb{N}$ such that $A_n \subseteq B_n$. From this,

$$\begin{aligned} 1 = d(\lambda) &\geq \sum_{w \in B_n} d(w)2^{-sn} \\ &\geq \sum_{w \in B_n} 2^{-sn} \\ &\geq \sum_{w \in A_n} 2^{-sn} \\ &= |A_n|2^{-sn} \end{aligned}$$

Therefore, $|A_n| < 2^{sn}$ and by (8), $2^{sn} > (1-\epsilon)2^{(h-\epsilon)n}$. Since this holds for all $\epsilon > 0$, it follows that $\dim_{\mathrm{H}}(\mathcal{T}(\mu)) \geq h$.

**Proof of Proposition 5.7.** Let $\alpha \in [0,1]$. We first prove that $\mathcal{H}(\mu^\alpha) = \mathcal{H}(\alpha)$. Notice that $\mu^\alpha$ can be defined as the product measure based on $\mu_1^\alpha : \{0,1\} \to [0,1]$ where $\mu_1^\alpha(b) = \mu^\alpha(b)$. That is,

$$\mu^\alpha(w) = \mu_1^\alpha(w[0])\mu_1^\alpha([2]) \ldots \mu_1^\alpha(w[n-1]) \tag{9}$$

for each $n \in \mathbb{N}$ and $w \in \{0,1\}^n$. This it follows immediately from

$$\mu_1^\alpha(b) = \begin{cases} \alpha & \text{if } b = 1 \\ 1-\alpha & \text{if } b = 0 \end{cases}$$

Then we find that

$$\mathcal{H}_n(\mu_n^\alpha) = \mathcal{H}_1(\mu_1^\alpha) + \mathcal{H}_1(\mu_1^\alpha) \ldots + \mathcal{H}_1(\mu_1^\alpha) = n\mathcal{H}_1(\mu_1^\alpha)$$

and from that

$$\mathcal{H}(\mu) = \lim_n \frac{n\mathcal{H}_1(\mu_1^\alpha)}{n} = \mu^\alpha(0)\log\frac{1}{\mu^\alpha(0)} + \mu^\alpha(1)\log\frac{1}{\mu^\alpha(1)} = \mathcal{H}(\alpha)$$

We let now prove that $\mu^\alpha$ is a stationary ergodic measure. That is, the measure induced by $\mu^\alpha$,

$$\begin{array}{rcl} \nu^\alpha : \mathcal{F} & \to & [0,1] \\ C_w & \mapsto & \mu^\alpha(w) \end{array}$$

is stationary and ergodic.

To see that $\nu^\alpha$ is stationary it suffices to prove that for all $w \in \{0,1\}^*$

$$\nu^\alpha(T^{-1}C_w) = \nu^\alpha(C_w). \quad (10)$$

This is because $\mathcal{F}$ is the $\sigma$-algebra generated from $\{C_w \mid w \in \{0,1\}^*\}$. Notice that $T^{-1}(C_w) = C_{0w} \cup C_{1w}$ and (10) is equivalent to

$$\mu^\alpha(0w) + \mu^\alpha(1w) = \mu^\alpha(w)$$

that it follows from 9.

We let prove that $\nu^\alpha$ is an ergodic measure. Let $w \in \{0,1\}^n$ and $u \in \{0,1\}^m$, we have that

$$\nu^\alpha(T^{-N}(C_w) \cap C_u) = \nu^\alpha(C_w)\nu^\alpha(C_u)$$

for sufficient large values of $N$ (for example $N > m$). Then, for all $X_1, X_2 \in \mathcal{F}$ we have that

$$\lim_{n\to\infty} \nu^\alpha(T^{-n}X_1 \cap X_2) = \nu^\alpha(X_1)\nu^\alpha(X_2)$$

This implies ergodicity, for if $T^{-1}X = X$ then we have that $\nu^\alpha(X) = \nu^\alpha(X \cap X) = \nu^\alpha(X)\nu^\alpha(X)$ and $\nu^\alpha(X) = 0$ or $\nu^\alpha(X) = 1$.



ECCC ISSN 1433-8092

http://eccc.hpi-web.de/