Multi-Head Finite-State Dimension

Xiang Huang
University of Illinois Springfield, IL, USA

Xiaoyuan Li
Iowa State University, Ames, IA, USA

Jack H. Lutz
Iowa State University, Ames, IA, USA

Neil Lutz
University of Notre Dame, IN, USA

Abstract

We introduce multi-head finite-state dimension, a generalization of finite-state dimension in which a group of finite-state agents (the heads) with oblivious, one-way movement rules, each reporting only one symbol at a time, enable their leader to bet on subsequent symbols in an infinite data stream. In aggregate, such a scheme constitutes an h-head finite state gambler whose maximum achievable growth rate of capital in this task, quantified using betting strategies called gales, determines the multi-head finite-state dimension of the sequence. The 1-head case is equivalent to finite-state dimension as defined by Dai, Lathrop, Lutz and Mayordomo (2004). In our main theorem, we prove a strict hierarchy as the number of heads increases, giving an explicit sequence family that separates, for each positive integer h, the earning power of h-head finite-state gamblers from that of $(h+1)$-head finite-state gamblers. We prove that multi-head finite-state dimension is stable under finite unions but that the corresponding quantity for any fixed number $h > 1$ of heads – the h-head finite-state predimension – lacks this stability property.

2012 ACM Subject Classification Theory of computation → Automata over infinite objects

Keywords and phrases Finite-state dimension, effective dimension, algorithmic randomness

Digital Object Identifier 10.4230/LIPIcs.MFCS.2026.59

Related Version *Full Version*: https://arxiv.org/abs/2509.22912

Funding *Xiang Huang*: This research was supported in part by U.S. Department of Energy Grant DE-SC0024278.
Xiaoyuan Li: This research was supported in part by National Science Foundation Grants 1545028 and 1900716.
Jack H. Lutz: This research was supported in part by National Science Foundation Grants 1545028 and 1900716.
Neil Lutz: This research was conducted while Neil Lutz was at Swarthmore College.

1 Introduction

This paper introduces *multi-head finite-state dimension* and *multi-head finite-state strong dimension*, two new tools for algorithmic information theory. Our main theorem is a hierarchy result establishing that each additional head strengthens these multi-head dimensions.

The primary historical steps leading to our work are as follows.

(i) Hausdorff [18] introduced *Hausdorff dimension*, which became a central tool in classical geometric measure theory. (Note the term "classical" here does not mean "old," but rather not involving the theory of computation or related aspects of mathematical logic.)

(ii) Tricot [47] and Sullivan [46] introduced *packing dimension*, a kind of dual and upper bound of Hausdorff dimension.

(iii) J. Lutz [30, 31] proved a *gale characterization* of Hausdorff dimension that enables *effectivizations* of Hausdorff dimension at various levels of computability and computational complexity.

(iv) Dai, Lathrop, J. Lutz, and Mayordomo [10] showed that Hausdorff dimension can even be effectivized at the finite-state level.

(v) Athreya, Hitchcock, J. Lutz, and Mayordomo [3] proved that the above effectivizations can also be carried out for packing dimension, thus yielding *strong dimensions* at each of these levels.

All the above have yielded many applications in algorithmic information theory, computational complexity, geometric measure theory, analytic number theory, and other areas. In particular, the finite-state dimension $\dim_{\mathrm{FS}}(S)$ and the finite-state strong dimension $\mathrm{Dim}_{\mathrm{FS}}(S)$ of infinite sequences S over a finite alphabet Σ have had useful interactions with Borel normal numbers and uniform distribution theory [6, 7, 11, 17, 27, 33, 29, 37, 38, 39, 40], Shannon entropy and Kullback-Leibler divergence [5, 21, 25, 26, 27, 32], and efficient prediction of data streams [13, 16, 20]. The references here are representative but far from exhaustive. The Effective Fractal Dimension Bibliography maintained by Hitchcock [19] is more comprehensive.

The present paper was motivated by a long-standing open question and inspired by an ingenious proof. The long-standing open question is whether the finite-state setting is the only level of effectivity in which randomness is equivalent to full dimension. Bourke, Hitchcock, and Vinodchandran [5] noted that the equivalence of finite-state randomness and finite-state full dimension follows from a theorem of Schnorr and Stimm [45]. In contrast, there is no such equivalence for algorithmic randomness and dimension, computable randomness and dimension, or resource-bounded randomness and dimension; in all these settings, there exist non-random sequences with full dimension.

The search for other settings where this equivalence holds has focused on notions of randomness and dimension that are "highly effective," defined using gamblers that are computationally weaker than polynomial-time algorithms. Examples include the settings of pushdown gamblers, Lempel-Ziv gamblers, and perhaps logspace gamblers [1, 2, 12, 29, 41, 35], but it is not yet known whether randomness and full dimension are equivalent in those settings. Here our objective is to investigate a setting that is as close as possible to, but not quite as effective as, the setting of finite-state gamblers.

The ingenious proof that inspired our choice of a setting to investigate is due to Becher, Carton, and Heiber [4]. Translated into the original gambling (i.e., gale) formulation of finite-state dimension [10], Becher, Carton, and Heiber proved that there is an infinite binary sequence S with the following two properties.

(1) No finite-state gambler G can win unbounded money betting on the successive bits of S.

(2) There is a gambler $\widehat{G}$ that wins money exponentially quickly betting on the successive bits of S. This gambler $\widehat{G}$ is "almost" finite-state, except that it is equipped with an additional *probe to the past*, so that at any time, $\widehat{G}$ knows not only the last bit of the string scanned so far, but also the "middle" bit of this string (adjusted appropriately for the fact that this string may have even or odd length).

Taken together, items (1) and (2) above imply that equipping a finite-state gambler with a "probe to the past" as above can make it more powerful than any finite-state gambler. This is intuitively unsurprising, because implementing such a probe requires access to an unbounded amount of past data.

In a line of research over sixty years old [43, 44], a device like $\widehat{G}$ is a *two-head finite automaton* whose first head is the "leading" (scanning) head of $\widehat{G}$ and whose second head is what we called above a "probe to the past." The *multi-head finite automata* of this literature (surveyed nicely by Holzer, Kutrib, and Malcher [24]) are typically used to recognize sets of finite strings rather than to compress or gamble on infinite data streams, but the present paper should be regarded as a continuation of this literature as well as a further development of algorithmic dimensions.

As noted above, the goal here is to investigate a setting that is as close as possible to, but not quite as effective as, the setting of finite-state gamblers. Accordingly, we require the trailing heads in our multi-head finite-state gamblers to be as weak as possible: They can only move forward (or hold still), and they are *oblivious* in the sense that their movements do not depend on the data that they are scanning. (We are using the terminology of oblivious Turing machines here [42]. In the terminology of [24], our trailing heads are *data-independent*.) Cruz, Glashausser, Li, and Lutz [8] recently studied a non-oblivious variant of our multi-head model and showed that, with only modest changes, the argument for our hierarchy theorem carries over into that setting. Cruz, Glashausser, and Lutz [9] further investigated the separation between these two models.

Section 3 below defines, for each positive integer h, the notion of an *h-head finite-state gambler* (*h-FSG*) in a straightforward way such that a 1-FSG is a finite-state gambler as in [10] and the gambler $\widehat{G}$ above is a 2-FSG. Section 4 then uses section 3 to define, for each positive integer h and each sequence $S \in \Sigma^\omega$, the *h-head finite-state predimension* $\dim_{\mathrm{FS}}^{(h)}(S)$ and the *h-head finite-state strong predimension* $\mathrm{Dim}_{\mathrm{FS}}^{(h)}(S)$ in such a way that $\dim_{\mathrm{FS}}^{(1)}(S) = \dim_{\mathrm{FS}}(S)$ is the finite-state dimension of S as in [10] and $\mathrm{Dim}_{\mathrm{FS}}^{(1)}(S) = \mathrm{Dim}_{\mathrm{FS}}(S)$ is the finite-state strong dimension of S as in [3]. It is immediately evident that each of these predimensions is, for fixed S, a nonincreasing function of h, leading immediately to the definitions of the *multi-head finite-state dimension*

$$\dim_{\mathrm{FS}}^{\mathrm{MH}}(S) = \inf_{h \in \mathbb{Z}^+} \dim_{\mathrm{FS}}^{(h)}(S)$$

and the *multi-head finite-state strong dimension*

$$\mathrm{Dim}_{\mathrm{FS}}^{\mathrm{MH}}(S) = \inf_{h \in \mathbb{Z}^+} \mathrm{Dim}_{\mathrm{FS}}^{(h)}(S).$$

We note that our model is definitionally robust, in that N. Lutz [34] has very recently shown that our multi-head finite-state predimensions and dimensions are equivalent to natural notions of multi-head finite-state *compressibility*.

We prove the main theorem of this paper, a hierarchy theorem, in section 5. Recalling that, for each $h \in \mathbb{Z}^+$ and each $S \in \Sigma^\omega$,

$$\dim_{\mathrm{FS}}^{(h+1)}(S) \le \dim_{\mathrm{FS}}^{(h)}(S) \quad \text{and} \quad \mathrm{Dim}_{\mathrm{FS}}^{(h+1)}(S) \le \mathrm{Dim}_{\mathrm{FS}}^{(h)}(S),$$

our main theorem says that each of these inequalities is strict for some sequences S. This hierarchy theorem is analogous to Yao and Rivest's nontrivial 1978 proof that, in the context of language recognition, "$k+1$ heads are better than k" [48] and the subsequent line of hierarchy theorems for the recognition power of multi-head finite automata in two-way [36, 23] or data-independent [23, 22, 14, 15] settings.

For each positive integer h, we define a function $F_{h+1} : \Sigma^\omega \to \Sigma^\omega$, where every sequence $F_{h+1}(S)$ has a structural pattern that is easily exploited by an $(h+1)$-FSG, but each of those $h+1$ heads is essential for this task. We prove that if R is a Martin-Löf random sequence, then no h-FSG can gain any significant advantage from the patterns in $F_{h+1}(R)$. Our proof relies on partitioning $F_{h+1}(R)$ into a family of strings of exponentially increasing length, where each string has high conditional Kolmogorov complexity given the information that can be accessed by any $(h-1)$ heads while the leading head reads that string.

Each fractal dimension $\mathtt{dim}$ (we use this font generically in this paragraph) assigns a dimension $\mathtt{dim}(E)$ to each *subset* E of some metric space, which is Σ^ω in this paper. If $\mathtt{dim}$ is one of the effective dimensions of algorithmic information theory, one then sets $\mathtt{dim}(S) = \mathtt{dim}(\{S\})$ for each $S \in \Sigma^\omega$, as we do here. (This would be senseless for *classical* fractal dimensions, because there all singletons have dimension 0.) In general, each fractal dimension $\mathtt{dim}$ has the *stability* property that, for all sets E and F,

$$\mathtt{dim}(E \cup F) = \max\{\mathtt{dim}(E), \mathtt{dim}(F)\}.$$

This distinguishes fractal dimensions, for example, from measures.

The finite-state dimensions $\dim_{\mathrm{FS}}$ and $\mathrm{Dim}_{\mathrm{FS}}$ are stable in the above sense. However, we prove in section 6 that, for each $h \geq 2$, the h-head finite-state predimensions $\dim_{\mathrm{FS}}^{(h)}$ and $\mathrm{Dim}_{\mathrm{FS}}^{(h)}$ are *not* stable. We prove in section 7 that the multi-head finite-state dimensions $\dim_{\mathrm{FS}}^{\mathrm{MH}}$ and $\mathrm{Dim}_{\mathrm{FS}}^{\mathrm{MH}}$ *are* stable. This is why we use the "predimension" terminology.

Some proofs are omitted in this conference version of the paper.

2 Preliminaries

String and Sequence Notation

Given a finite alphabet Σ of symbols, the sequence space Σ^ω is the set of all infinite sequences over Σ; when $\Sigma = \{0, 1\}$, this is the Cantor space $\mathbf{C}$. For $i \in \mathbb{N}$, we write $S[i]$ for the symbol at position i in a sequence $S \in \Sigma^\omega$; the leftmost symbol of S is $S[0]$. For $i, j \in \mathbb{N}$ with $i \leq j$, we write $S[i..j]$ for the string $S[i] \cdots S[j] \in \Sigma^{j-i+1}$.

s-Gales and Martingales

Following the standard template for effectivizations of fractal dimensions [30, 31, 3], we will base our dimension notions on the success of functions called gales.

▶ **Definition 2.1.** *For $s \in [0, \infty)$, an s-*gale *is a function $d : \Sigma^* \to [0, \infty)$ that satisfies the condition*

$$d(w) = |\Sigma|^{-s} \sum_{b \in \Sigma} d(wb)$$

for all $w \in \Sigma^*$. A martingale *is a 1-gale.*

Intuitively, the value of a gale is interpreted as the capital of a gambler betting on successive symbols in a sequence. The number s parameterizes the favorability of the betting environment. A martingale, where $d(w)$ is the average, over all symbols b, of $d(wb)$, corresponds to a perfectly fair betting environment, in which a maximally conservative gambler who places equal bets on all possible one-symbol extensions of w can maintain constant capital. Parameters $s < 1$ correspond to less favorable betting environments where "the house takes a cut." In these environments, the gambler would need to exploit some knowledge about the sequence – or be lucky – to maintain constant capital. For a gale to *succeed*, though, we require that the gambler's capital is unbounded, not merely constant.

▶ **Definition 2.2.** *Let d be an s-gale.*

1. *We say that d* succeeds *on a sequence $S \in \Sigma^\omega$ if $\limsup_{n\to\infty} d(S[0..n-1]) = \infty$. The* success set *of d is $S^\infty[d] = \{S \in \Sigma^\omega \mid d \text{ succeeds on } S\}$.*
2. *We say that d* succeeds strongly *on a sequence $S \in \Sigma^\omega$ if $\liminf_{n\to\infty} d(S[0..n-1]) = \infty$. The* strong success set *of d is $S^\infty_{\mathrm{str}}[d] = \{S \in \Sigma^\omega \mid d \text{ succeeds strongly on } S\}$.*

In general, succeeding on a sequence becomes more difficult as the sequence becomes more complex, as the computational resources available to the gale become more restricted, and as the value s decreases.

Kolmogorov Complexity

While the success of gales quantifies the predictability of sequences, Kolmogorov complexity quantifies the compressibility of strings. These are closely related concepts, and several of our proofs rely on bounding the compressibility of parts of a sequence to derive bounds on the sequence's overall predictability. Our proofs will use the basic properties of Kolmogorov complexity stated below. Further details and background on Kolmogorov complexity can be found in [28].

Let U be a fixed, universal prefix Turing machine. For strings $\sigma, \tau \in \{0,1\}^*$, the (prefix) *conditional Kolmogorov complexity* of σ given τ is

$$K(\sigma \mid \tau) = \min\{|\pi| \mid \pi \in \{0,1\}^* \text{ and } U(\pi, \tau) = \sigma\},$$

that is, the minimum length of a *program* π for σ given τ as an auxiliary input. The *Kolmogorov complexity* of σ is $K(\sigma) = K(\sigma \mid \lambda)$, where λ denotes the empty string.

- *Computable functions do not add information:* For every computable function f and all $\sigma, \tau \in \{0,1\}^*$,

$$K(f(\sigma) \mid \tau) \leq K(\sigma \mid \tau) + O(1). \tag{1}$$

- *Symmetry of information:* For all $\sigma, \tau \in \{0,1\}^*$,

$$K(\sigma, \tau) = K(\sigma \mid \tau, K(\tau)) + K(\tau) + O(1). \tag{2}$$

A sequence $R \in \mathbf{C}$ is *Martin-Löf random* if there is some constant $c \in \mathbb{N}$ such that, for all $n \in \mathbb{N}$, we have $K(R[0..n-1]) \geq n - c$; in this sense, such sequences are *incompressible*. We will construct sequences based on Martin-Löf random sequences, and our constructed sequences will inherit some of this incompressibility.

3 Multi-Head Finite-State Gamblers

We now define multi-head finite-state gamblers. These are a generalization of the 1-account finite-state gamblers defined in [10]; our definition here is self-contained. We first define the formal components of a gambler, then describe its movements and operation.

The gambler's state space will be the Cartesian product of two sets, T and Q. The T component of the state will govern the heads' movements, while the Q component will govern its betting behavior. We will place a requirement on the transition function δ that it can be decomposed into separate transition functions δ_T and δ_Q that act on T and Q, respectively, where δ_T does not depend on the symbols read. This has the effect of isolating the T component of the gambler's state from the tape contents, which makes the head movements data-independent.

▶ **Definition 3.1.** *Let $h \in \mathbb{Z}^+$. An h-*head finite-state gambler *(h-*FSG*) is a 7-tuple*

$$G = (T \times Q, \Sigma, \delta, \mu, \beta, (t_0, q_0), c_0),$$

where

- *T and Q are nonempty finite sets, and $T \times Q$ is the set of* states*;*
- *Σ is the finite alphabet, satisfying $|\Sigma| \geq 2$;*

- $\delta : T \times Q \times \Sigma^h \to T \times Q$ *is the* transition function*, with the following property: there are functions* $\delta_T : T \to T$ *and* $\delta_Q : Q \times \Sigma^h \to Q$ *such that, for all* $t \in T$*,* $q \in Q$*, and* $\vec{\sigma} \in \Sigma^h$*,*

$$\delta(t, q, \vec{\sigma}) = (\delta_T(t), \delta_Q(q, \vec{\sigma}));$$

- $\mu : T \to \{0,1\}^{h-1}$ *is the* movement function*;*
- $\beta : Q \to \Delta_{\mathbb{Q}}(\Sigma)$ *is the* betting function*, where* $\Delta_{\mathbb{Q}}(\Sigma)$ *denotes all rational-valued discrete probability distributions on* Σ*;*
- $(t_0, q_0) \in T \times Q$ *is the* initial state*; and*
- c_0 *is the* initial capital*.*

Head Movements and State Transitions

Intuitively, we regard G as having h *heads*: a *leading head* which advances through the input sequence in discrete steps, and $h-1$ *trailing heads*, some subset of which advance in each step. The positions of the heads are governed by the *positional states* – the members of T – and the function μ, which dictates, in each positional state, which of the trailing heads advance. More formally, for $r \in \mathbb{N}$, let $\delta_T^r : T \to T$ denote r iterated applications of δ_T. Then the *positional vector* of G after n steps, which tells us the positions of all trailing heads, is determined by the function $\pi : \mathbb{N} \to \mathbb{N}^{h-1}$ defined by $\pi(0) = (0, \ldots, 0)$ and, for all $n \in \mathbb{N}$, recursively by $\pi(n+1) = \pi(n) + \mu\left(\delta_T^n(t_0)\right)$.

For $n \in \mathbb{N}$ and $S \in \Sigma^\omega$, let $(\pi_1, \ldots, \pi_{h-1}) = \pi(n)$ and $\vec{\sigma}_n = (S[\pi_1], \ldots, S[\pi_{h-1}], S[n])$. Then the state sequence $(t_0, q_0), (t_1, q_1), (t_2, q_2), \ldots$ of G on input sequence $S \in \Sigma^\omega$ is given, for all $n \in \mathbb{N}$, by

$$(t_{n+1}, q_{n+1}) = \delta(t_n, q_n, \vec{\sigma}_n). \tag{3}$$

Note that the head movements are *oblivious* in the sense of being data-independent, which, together with the finite-state condition, implies that each of them moves at an essentially constant rate, in the following sense.

▶ **Observation 3.2.** *For each trailing head* i*, there is a constant* $\sigma_i \in [0,1]$*, which we call the* speed *of head* i*, such that whenever the leading head is at position* $n \in \mathbb{N}$*, the position of head* i *is between* $\sigma_i n - |T|$ *and* $\sigma_i n + |T|$*, inclusive.*

Proof. As T is finite, repeated application of δ_T to t_0 must eventually result in a cycle, which will thereafter be repeated forever. Let L be the set of positional states in this cycle.

Let D be the set of positional states that appeared before entering the cycle. Note that D and L are disjoint. After n steps, the cycle has been repeated $(n - |D|)/|L|$ times. For each trailing head i, let $\sigma_i \in [0,1]$ be such that head i advances $\sigma_i |L|$ times during the cycle. Then the position of head i after n steps is bounded below by $\sigma_i |L| \left\lfloor \frac{n-|D|}{|L|} \right\rfloor \geq \sigma_i n - |D| - |L| \geq \sigma_i n - |T|$ and above by $|D| + \sigma_i |L| \left\lceil \frac{n-|D|}{|L|} \right\rceil \leq \sigma_i n + |D| + |L| \leq \sigma_i n + |T|$. ◀

4 Multi-Head Finite-State Dimensions

We now define h-head finite-state (strong) predimension – which generalize the finite-state dimension of [10] and the finite-state strong dimension of [3] – and multi-head finite-state (strong) dimension. Definitions 4.1–4.4 continue to follow the standard template for effectivizing fractal dimensions [30, 31, 3], differing from the corresponding definitions in [10, 3] only in that the gamblers now have h heads.

Briefly, the approach is to define the martingale and s-gales induced by the betting function of a specific gambler, then consider the range of parameters s for which it is possible for a gambler's s-gale to (strongly) succeed on a given set. Recalling that s parameterizes the favorability of the betting environment, this means we are quantifying how unfavorable the betting environment can become before it becomes impossible for an h-FSG to succeed on the set. Intuitively, a gambler can succeed in a more unfavorable environment if the set only contains simple, predictable sequences, so such sets will tend to have lower predimension or dimension.

▶ **Definition 4.1.** *The martingale of an h-FSG G is the function $d_G : \Sigma^* \to [0, \infty)$ defined recursively by $d_G(\lambda) = c_0$ and $d_G(wb) = |\Sigma| d_G(w)\beta(q_{|w|})(b)$, where $q_{|w|}$ is defined as in* (3)*.*

By rescaling this martingale, we derive corresponding s-gales.

▶ **Definition 4.2.** *For $s \in [0, \infty)$, the s-gale of an h-FSG G is the function $d_G^{(s)} : \Sigma^* \to [0, \infty)$ defined, for all $w \in \Sigma^*$, by $d_G^{(s)}(w) = |\Sigma|^{(s-1)|w|} d_G(w)$.*

Given a set of sequences, we may then consider the range of parameters s for which h-FSGs with (strongly) successful s-gales exist.

▶ **Definition 4.3.** *Let $X \subseteq \Sigma^\omega$.*
1. $\mathcal{G}_h(X) = \left\{ s \in [0, \infty) \;\middle|\; \exists h\text{-}FSG\ G \text{ with } X \subseteq S^\infty\big[d_G^{(s)}\big] \right\}$*.*
2. $\mathcal{G}_h^{\mathrm{str}}(X) = \left\{ s \in [0, \infty) \;\middle|\; \exists h\text{-}FSG\ G \text{ with } X \subseteq S^\infty_{\mathrm{str}}\big[d_G^{(s)}\big] \right\}$*.*

The infima of these ranges then give us our predimensions; the h-head finite-state (strong) predimension of a set X is the infimum of the parameters s such that an h-FSG can have an s-gale that (strongly) succeeds on X.

▶ **Definition 4.4.** *Let $h \in \mathbb{Z}^+$ and $X \subseteq \Sigma^\omega$.*
1. *The* h-head finite-state predimension *of X is* $\dim_{\mathrm{FS}}^{(h)}(X) = \inf \mathcal{G}_h(X)$*.*
2. *The* h-head finite-state strong predimension *of X is* $\mathrm{Dim}_{\mathrm{FS}}^{(h)}(X) = \inf \mathcal{G}_h^{\mathrm{str}}(X)$*.*

Note that $\dim_{\mathrm{FS}}^{(1)}$ is the finite-state dimension $\dim_{\mathrm{FS}}$ defined in [10] and $\mathrm{Dim}_{\mathrm{FS}}^{(1)}$ is the finite-state strong dimension $\mathrm{Dim}_{\mathrm{FS}}$ defined in [3].

▶ **Definition 4.5.** *Let $X \subseteq \Sigma^\omega$.*
1. *The* multi-head finite-state dimension *of X is* $\dim_{\mathrm{FS}}^{\mathrm{MH}}(X) = \inf_{h \in \mathbb{Z}^+} \dim_{\mathrm{FS}}^{(h)}(X)$*.*
2. *The* multi-head finite-state strong dimension *of X is* $\mathrm{Dim}_{\mathrm{FS}}^{\mathrm{MH}}(X) = \inf_{h \in \mathbb{Z}^+} \mathrm{Dim}_{\mathrm{FS}}^{(h)}(X)$*.*

Both h-head finite-state predimension and multi-head finite-state dimension are monotone, in the following sense.

▶ **Observation 4.6.** *Let $X, Y \subseteq \Sigma^\omega$.*
1. *For all $h \in \mathbb{Z}^+$, if $X \subseteq Y$, then $\dim_{\mathrm{FS}}^{(h)}(X) \leq \dim_{\mathrm{FS}}^{(h)}(Y)$ and $\mathrm{Dim}_{\mathrm{FS}}^{(h)}(X) \leq \mathrm{Dim}_{\mathrm{FS}}^{(h)}(Y)$.*
2. *If $X \subseteq Y$, then $\dim_{\mathrm{FS}}^{\mathrm{MH}}(X) \leq \dim_{\mathrm{FS}}^{\mathrm{MH}}(Y)$ and $\mathrm{Dim}_{\mathrm{FS}}^{\mathrm{MH}}(X) \leq \mathrm{Dim}_{\mathrm{FS}}^{\mathrm{MH}}(Y)$.*

Proof. Let $h \in \mathbb{Z}^+$ and assume $X \subseteq Y$. By Definition 4.3, for each $s \in \mathcal{G}_h(Y)$ there is some h-FSG G such that $X \subseteq Y \subseteq S^\infty\big[d_G^{(s)}\big]$, so $s \in \mathcal{G}_h(X)$. Hence, $\mathcal{G}_h(Y) \subseteq \mathcal{G}_h(X)$, and therefore $\dim_{\mathrm{FS}}^{(h)}(X) \leq \dim_{\mathrm{FS}}^{(h)}(Y)$, by Definition 4.4. $\mathrm{Dim}_{\mathrm{FS}}^{(h)}(X) \leq \mathrm{Dim}_{\mathrm{FS}}^{(h)}(Y)$ holds by the same argument.

By Definition 4.5, the second statement follows immediately from the first. ◀

▶ **Definition 4.7.** *The* h-head finite-state predimension, h-head finite-state strong predimension, multi-head finite-state dimension, *and* multi-head finite-state strong dimension *of an individual sequence $S \in \Sigma^\omega$ are the h-head finite-state predimension, h-head finite-state strong predimension, multi-head finite-state dimension, and multi-head finite-state strong dimension, respectively, of the singleton $\{S\}$.*

5 A Hierarchy of Predimensions

In this section, we prove our main theorem, a strict hierarchy on predimensions, showing that additional heads enhance the predictive power of multi-head finite-state gamblers.

▶ **Theorem 5.1.** *For each $h \in \mathbb{Z}^+$, there is a sequence Y such that*

$$\dim_{\mathrm{FS}}^{(h+1)}(Y) < \dim_{\mathrm{FS}}^{(h)}(Y) \quad \text{and} \quad \mathrm{Dim}_{\mathrm{FS}}^{(h+1)}(Y) < \mathrm{Dim}_{\mathrm{FS}}^{(h)}(Y).$$

We will prove this by defining an explicit family of functions on sequences, which we will then apply to Martin-Löf random sequences. The goal, for each $h \in \mathbb{Z}^+$, is to define sequences in which a constant fraction of the bits depend on exactly h well-spaced bits from earlier in the sequence, so that h trailing heads are necessary and sufficient to access the relevant earlier bits.

▶ **Definition 5.2.** *For each $i \in \mathbb{Z}^+$, let p_i denote the i^{th} prime number: $p_1 = 2$, $p_2 = 3$, etc. For each $x \in \mathbb{N}$, define the* multiplicity *of p_i in x to be $\nu_i(x) = \max\left\{k \in \mathbb{N} \;\middle|\; p_i^k \mid x\right\}$.*

▶ **Definition 5.3.** *For each $h \in \mathbb{Z}^+$, define the function $F_{h+1} : \mathbf{C} \to \mathbf{C}$ as follows. For all $S \in \mathbf{C}$, let $F_{h+1}(S)[0] = S[0]$ and, for all $q \in \mathbb{N}$ and $0 \le r < p_{h+1}$, let*

$$F_{h+1}(S)[qp_{h+1} + r] = \begin{cases} S[q(p_{h+1} - 1) + r] & \text{if } r > 0 \\ \bigoplus_{k=1}^{h} F_{h+1}(S)[qp_k] & \text{if } r = 0, \end{cases}$$

where $\oplus$ is the parity operation.

For example, in the sequence $F_3(S)$, four out of every five bits are "new" bits taken directly from S, and the fifth bit, at some index $5q$, is given by the parity of the bits at indices $2q$ and $3q$ in $F_3(S)$. Those bits themselves might come directly from S or, if q is a multiple of 5, the recursive dependence may continue. Note that an index in S may be reachable by an even number of paths in the recursion tree whose root is at $5q$, in which case cancellation makes $F_3(S)[5q]$ insensitive to the bit at that index. In Figure 1 we show an example of the depth-2 recursive calculation of $F_3(S)[150]$, in which $S[29]$ cancels itself out.

▶ **Lemma 5.4.** *For all $S \in \mathbf{C}$ and $h \in \mathbb{Z}^+$,*

$$\dim_{\mathrm{FS}}^{(h+1)}(F_{h+1}(S)) \le \mathrm{Dim}_{\mathrm{FS}}^{(h+1)}(F_{h+1}(S)) \le 1 - \frac{1}{p_{h+1}}.$$

Proof sketch. For each $h \in \mathbb{Z}^+$, we construct an $(h+1)$-head finite-state gambler G such that, for all $\epsilon > 0$ and all $S \in \mathbf{C}$, the sequence $F_{h+1}(S)$ is in the success set of the $(1 - \frac{1}{p_{h+1}} + \epsilon)$-gale of G. In each round of p_{h+1} steps, the i^{th} head of G advances p_i times, such that the trailing heads point to the bits referenced in the next parity calculation. This means that G can predict one out of every p_{h+1} bits with certainty, and it bets uniformly on all other bits. This corresponds to its $(1 - \frac{1}{p_{h+1}} + \epsilon)$-gale amassing unbounded capital. ◀

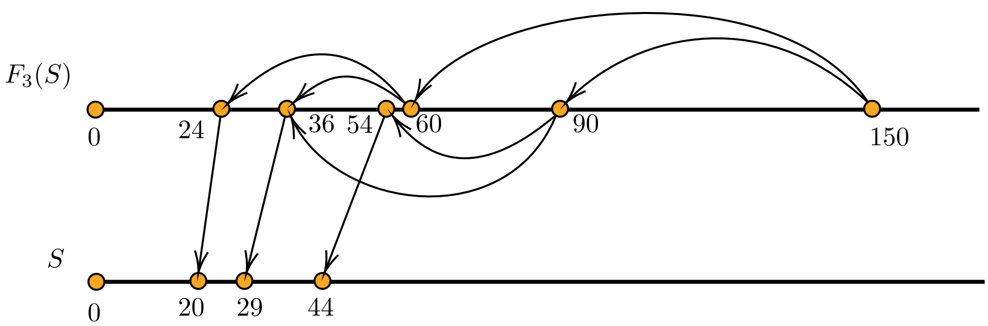

Figure 1 $F_3(S)[150] = S[20] \oplus S[29] \oplus S[29] \oplus S[44] = S[20] \oplus S[44]$.

Fix any positive integer h and any Martin-Löf random sequence $R \in \mathbf{C}$, and let $Y = F_{h+1}(R)$. To prove Theorem 5.1, it suffices to show that the h-head finite-state predimension of Y is 1. To do this, our high-level plan is to show that for every h-FSG and each sufficiently long prefix $Y[0..n]$ of Y, there is a long suffix $Y[b..n]$ that has high conditional Kolmogorov complexity given the information the trailing heads might read while the leading head reads $Y[b..n]$. This high conditional complexity will bound the rate of growth of the h-FSG's martingale on that suffix, and we will apply this algorithmic information-theoretic argument iteratively to bound the martingale's overall rate of growth.

Let G be any h-FSG over the alphabet $\{0,1\}$. Let $\sigma_1, \dots, \sigma_{h-1}$ be the trailing head speeds guaranteed by Observation 3.2, and let $c = |T|$. Without loss of generality, assume $0 < \sigma_1 < \ldots < \sigma_{h-1} < 1$.

We now introduce some additional notation that will be helpful in carefully analyzing the recursive structure of the sequence Y. Let $n \in \mathbb{N}$. For every set $A \subseteq [0, n]$, define the string $Y[A] \in \{0,1\}^{n+1}$ by

$$Y[A][i] = \begin{cases} Y[i] & \text{if } i \in A \\ 0 & \text{otherwise,} \end{cases}$$

for all $i \in \{0, \dots, n\}$. Note that $Y[A] = Y[A \cap \mathbb{N}]$; we will only use integer indices of strings and sequences, but it will sometimes be convenient to consider the real intervals those integers belong to.

Let $b = \lfloor \gamma n \rfloor + 1$, where $\gamma \in (0,1) \cap \mathbb{Q}$ is a parameter whose value will be defined in Lemma 5.5. Define $u_h = Y[b..n]$,

$$U = \bigcup_{i=1}^{h-1} [\sigma_i b - c, \sigma_i n + c],$$

and $u = Y[U]$. Intuitively, the string u contains all bits that might be read by any trailing head while the leading head h reads u_h.

Let $d \in \mathbb{N}$ be a constant that does not depend on n. For each $1 \le i \le h$, define

$$V_i = \left\{ (p_i/p_{h+1})^{\nu_{h+1}(m)} m \;\middle|\; m \in [b..n] \text{ and } \nu_{h+1}(m) \le d \right\}$$

and its *closure*

$$\overline{V}_i = \left\{ m \cdot \frac{p_{h+1}^{a_1 + \ldots + a_h}}{p_1^{a_1} \cdots p_h^{a_h}} \;\middle|\; m \in V_i \text{ and } a_1, \dots, a_h \in \mathbb{N} \right\}.$$

Note that $V_i = \overline{V}_i \cap \{m \mid \nu_{h+1}(m) = 0\}$. Define $v_i = Y[V_i]$ and $\overline{v}_i = Y[\overline{V}_i]$.

To prove Theorem 5.1, we use the following four lemmas. Informally, Lemma 5.5 shows that we can tune parameters such that u excludes one of the h regions relevant to calculating parity bits in the suffix u_h; Lemma 5.6 shows the algorithmic information consequence that u_h has high complexity conditioned on u; Lemma 5.7 shows that when this conditional Kolmogorov complexity is high, a gambler cannot win quickly while betting on the suffix u_h; and Lemma 5.8 applies this reasoning inductively to show that a gambler cannot win much on any prefix of Y.

▶ **Lemma 5.5.** *There are constants (depending on c, d, h, and the trailing head speeds, but not on n) $\gamma \in (1/2, 1) \cap \mathbb{Q}$, $n_0 \in \mathbb{N}$, and $1 \leq j \leq h$ such that if $n \geq n_0$, then $U \cap \overline{V}_j = \emptyset$.*

▶ **Lemma 5.6.** *Fix γ and j as in Lemma 5.5, let $n \in \mathbb{N}$ be sufficiently large, and let $b = \lfloor \gamma n \rfloor + 1$. Then $K(u_h \mid u) \geq (n - b)(1 - p_{h+1}^{-d-1}) - O(\log n)$.*

▶ **Lemma 5.7.** *Let $h \in \mathbb{Z}^+$, G an h-FSG, $\varepsilon > 0$ a rational constant, and $n \in \mathbb{N}$ sufficiently large. Define b, u, and u_h as above. Let $y = Y[0..b-1]$. If $K(u_h \mid u) \geq (n-b)(1 - \frac{\varepsilon}{2})$, then*

$$\max_{w \sqsubseteq u_h} \frac{d_G(yw)}{d_G(y)} \leq 2^{\varepsilon \cdot (n-b)}. \tag{4}$$

▶ **Lemma 5.8.** *Let $Y = F_{h+1}(R)$, where R is a Martin-Löf random sequence and $h \in \mathbb{Z}^+$. For all $\epsilon \in \mathbb{Q} \cap (0, 1)$ and every h-head finite-state gambler G, the set $\left\{ d_G^{(1-2\epsilon)}(Y[0..n-1]) \;\middle|\; n \in \mathbb{N} \right\}$ is bounded by a constant.*

With these lemmas, we prove Theorem 5.1. Letting ϵ approach 0 and applying Lemma 5.8 shows that $\dim_{\mathrm{FS}}^{(h)}(Y) = \mathrm{Dim}_{\mathrm{FS}}^{(h)}(Y) = 1$, while Lemma 5.4 shows that

$$\dim_{\mathrm{FS}}^{(h+1)}(Y) \leq \mathrm{Dim}_{\mathrm{FS}}^{(h+1)}(Y) \leq 1 - \frac{1}{p_{h+1}}.$$

Hence, for every $h \in \mathbb{Z}^+$, $Y = F_{h+1}(R)$ attests to a strict separation between h-head and $(h+1)$-head finite-state predimension, and between h-head and $(h+1)$-head finite-state strong predimension.

6 Instability of h-Head Finite-State Predimensions

We now show that unlike typical fractal dimension notions, h-head finite-state predimensions are not stable for fixed $h \geq 2$.

▶ **Theorem 6.1.** *For each integer $h \geq 2$, h-head finite-state predimension and strong predimension are not stable under finite unions.*

We prove this using two variants of our function family F_{h+1}, called F'_{h+1} and F''_{h+1}. Recall that the bit in one out of every $h+1$ positions in a sequence $F_{h+1}(S)$ can be computed as the parity of h previous bits in the sequence. At those same positions in $F'_{h+1}(S)$ and $F''_{h+1}(S)$, the bit can be computed as the parity of only $h-1$ of the h bits involved in the $F_{h+1}(S)$ calculation. In $F'_{h+1}(S)$, it is parity of the first $h-1$ of those h bits, while in $F''_{h+1}(S)$, it is the parity of the last $h-1$ of those h bits. Intuitively, these sequences together capture the full complexity of $F_{h+1}(S)$, but each of them individually is well-suited to using only $h-1$ trailing heads.

Formally, for each $h \in \mathbb{Z}^+$, define the functions $F'_{h+1} : \mathbf{C} \to \mathbf{C}$ and $F''_{h+1} : \mathbf{C} \to \mathbf{C}$ as follows. For all $S \in \mathbf{C}$, let $F'_{h+1}(S)[0] = F''_{h+1}(S)[0] = S[0]$ and, for all $q \in \mathbb{N}$ and $0 \le r < p_{h+1}$, let

$$F'_{h+1}(S)[qp_{h+1} + r] = \begin{cases} S[q(p_{h+1} - 1) + r] & \text{if } r > 0 \\ \bigoplus_{k=1}^{h-1} F'_{h+1}(S)[qp_k] & \text{if } r = 0 \end{cases}$$

and

$$F''_{h+1}(S)[qp_{h+1} + r] = \begin{cases} S[q(p_{h+1} - 1) + r] & \text{if } r > 0 \\ \bigoplus_{k=2}^{h} F''_{h+1}(S)[qp_k] & \text{if } r = 0. \end{cases}$$

For any fixed h and $\epsilon > 0$, we can construct a pair of h-FSGs whose $(1 - \frac{1}{p_{h+1}} + \epsilon)$-gales succeed, for all $S \in \mathbf{C}$, on $F'_{h+1}(S)$ and on $F''_{h+1}(S)$, respectively, using constructions almost identical to the $(h+1)$-FSG that succeeds on $F_{h+1}(S)$ in Lemma 5.4; essentially, we omit either the first or the last trailing head from that construction. This establishes, for all $S \in \mathbf{C}$, that the singletons $\{F'_{h+1}(S)\}$ and $\{F''_{h+1}(S)\}$ have both h-head finite-state predimension and strong predimension at most $1 - \frac{1}{p_{h+1}}$.

Nevertheless, we show – using techniques very similar to those of our strict hierarchy result, Theorem 5.1 – that no h-FSG has a $(1-2\epsilon)$-gale that succeeds on *both* $F'_{h+1}(R)$ and $F''_{h+1}(R)$ when R is any Martin-Löf random sequence. This implies that the set $\{F'_{h+1}(R), F''_{h+1}(R)\}$ has h-head finite-state predimension and strong predimension 1, proving instability.

At a high level, the argument is that by the obliviousness of the head movements, the heads of any h-FSG must move identically on both these sequences, which means that for at least one of these sequences, at least one of the relevant $h - 1$ bits in any parity calculation must be missing from the information the gambler can readily access. Much as in Lemmas 5.6, 5.7, and 5.8, we show that this missing information implies that at least one of the sequences can be broken into strings of high conditional Kolmogorov complexity given the trailing heads' information, which bounds the corresponding gale's rate of growth on each of these strings and thereby on the entire sequence.

7 Stability of Multi-Head Finite-State Dimension

In this section we prove that, notwithstanding the instability of h-head finite-state predimension and strong predimension, multi-head finite-state dimension and strong dimension are stable under finite unions. This is because, as we now prove, a multi-head finite-state gambler with sufficiently many heads can implement a martingale that *approximately averages* the martingales of a pair of other multi-head finite-state gamblers, thereby succeeding wherever either of the pair would succeed, up to ϵ error. Maintaining this approximate average with finitely many states is delicate.

▶ **Theorem 7.1.** *Given $h_1, h_2 \in \mathbb{Z}^+$, for each h_1-FSG G_1, each h_2-FSG G_2 and each $\epsilon \in (0, 1)$, there is a $(h_1 + h_2 - 1)$-FSG G such that for all $s \in [0, 1]$,*

$$S^\infty\big[d_{G_1}^{(s)}\big] \cup S^\infty\big[d_{G_2}^{(s)}\big] \subseteq S^\infty\big[d_G^{(s+\epsilon)}\big] \quad \textit{and} \quad S^\infty_{\text{str}}\big[d_{G_1}^{(s)}\big] \cup S^\infty_{\text{str}}\big[d_{G_2}^{(s)}\big] \subseteq S^\infty_{\text{str}}\big[d_G^{(s+\epsilon)}\big].$$

▶ **Corollary 7.2.** *For all $X, Y \subseteq \Sigma^\omega$,*

$$\dim_{\text{FS}}^{\text{MH}}(X \cup Y) = \max\{\dim_{\text{FS}}^{\text{MH}}(X), \dim_{\text{FS}}^{\text{MH}}(Y)\}$$

and

$$\text{Dim}_{\text{FS}}^{\text{MH}}(X \cup Y) = \max\{\text{Dim}_{\text{FS}}^{\text{MH}}(X), \text{Dim}_{\text{FS}}^{\text{MH}}(Y)\}.$$

Proof. It is immediate from Theorem 7.1 that $\dim_{\mathrm{FS}}^{\mathrm{MH}}(X \cup Y) \le \max\{\dim_{\mathrm{FS}}^{\mathrm{MH}}(X), \dim_{\mathrm{FS}}^{\mathrm{MH}}(Y)\}$ and $\mathrm{Dim}_{\mathrm{FS}}^{\mathrm{MH}}(X \cup Y) \le \max\{\mathrm{Dim}_{\mathrm{FS}}^{\mathrm{MH}}(X), \mathrm{Dim}_{\mathrm{FS}}^{\mathrm{MH}}(Y)\}$. The reverse inequalities follow from Observation 4.6. ◀

8 Conclusion

To our knowledge, the finite-state dimensions $\dim_{\mathrm{FS}}$ and $\mathrm{Dim}_{\mathrm{FS}}$ are the most effective (meaning effectivized in the most restrictive way) of the fractal dimensions that have been investigated. It is not hard to see that the multi-head finite-state dimensions $\dim_{\mathrm{FS}}^{\mathrm{MH}}$ and $\mathrm{Dim}_{\mathrm{FS}}^{\mathrm{MH}}$, while not as effective as the finite-state dimensions, are *highly effective* in the sense that they are more effective than the polynomial-time dimensions $\dim_{\mathrm{p}}$ and $\mathrm{Dim}_{\mathrm{p}}$ of [30, 3].

Other dimensions that are highly effective in the above sense have been investigated. These include pushdown dimensions, Lempel–Ziv dimensions, and perhaps polylogspace dimensions [1, 2, 12, 29, 41, 35]. It will be interesting to see how multi-head finite-state dimensions are related to these other highly effective dimensions.

It will be noted that our motivating question remains open: A sequence $S \in \Sigma^\omega$ is *multi-head finite-state random* if there is no multi-head finite-state gambler G such that $S \in S^\infty[d_G]$. Is every sequence with multi-head finite-state dimension 1 multi-head finite-state random?

References

1 Pilar Albert, Elvira Mayordomo, and Philippe Moser. Bounded pushdown dimension vs Lempel Ziv information density. In Adam Day, Michael Fellows, Noam Greenberg, Bakhadyr Khoussainov, Alexander Melnikov, and Frances Rosamond, editors, *Computability and Complexity: Essays Dedicated to Rodney G. Downey on the Occasion of His 60th Birthday*, pages 95–114. Springer, Cham, 2017. doi:10.1007/978-3-319-50062-1_7.

2 Pilar Albert, Elvira Mayordomo, Philippe Moser, and Sylvain Perifel. Pushdown compression. In Susanne Albers and Pascal Weil, editors, *25th International Symposium on Theoretical Aspects of Computer Science*, volume 1 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 39–48, Dagstuhl, Germany, 2008. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.STACS.2008.1332.

3 Krishna B. Athreya, John M. Hitchcock, Jack H. Lutz, and Elvira Mayordomo. Effective strong dimension in algorithmic information and computational complexity. *SIAM Journal on Computing*, 37(3):671–705, 2007. doi:10.1137/S0097539703446912.

4 Verónica Becher, Olivier Carton, and Pablo A. Heiber. Finite-state independence. *Theory of Computing Systems*, 62:1555–1572, 2018. doi:10.1007/s00224-017-9821-6.

5 Chris Bourke, John M. Hitchcock, and N. V. Vinodchandran. Entropy rates and finite-state dimension. *Theoretical Computer Science*, 349(3):392–406, 2005. doi:10.1016/j.tcs.2005.09.040.

6 Yann Bugeaud. *Distribution Modulo One and Diophantine Approximation*. Cambridge Tracts in Mathematics. Cambridge University Press, 2012. doi:10.1017/CBO9781139017732.

7 Joe Clanin and Matthew Rayman. Finite-state dimension and the Davenport Erdős theorem. In *Mathematical Foundations of Computer Science 2026*, Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026. doi:10.4230/LIPIcs.MFCS.2026.38.

8 Julianne Cruz, Sho Glashausser, Xiaoyuan Li, and Neil Lutz. Adaptive multi-head finite-state gamblers. In *Computability in Europe (CiE 2026): Timeless Machines/Computability Across Eras*, Lecture Notes in Computer Science. Springer, to appear.

9 Julianne Cruz, Sho Glashausser, and Neil Lutz. One adaptive trailing head can outperform many oblivious trailing heads. In *Machines, Computations, and Universality (MCU 2026)*, Lecture Notes in Computer Science. Springer, to appear.

10 Jack J. Dai, James I. Lathrop, Jack H. Lutz, and Elvira Mayordomo. Finite-state dimension. *Theoretical Computer Science*, 310(1):1–33, 2004. `doi:10.1016/S0304-3975(03)00244-5`.

11 David Doty, Jack H. Lutz, and Satyadev Nandakumar. Finite-state dimension and real arithmetic. *Information and Computation*, 205(11):1640–1651, 2007. `doi:10.1016/j.ic.2007.05.003`.

12 David Doty and Jared Nichols. Pushdown dimension. *Theoretical Computer Science*, 381(1):105–123, 2007. `doi:10.1016/j.tcs.2007.04.005`.

13 Andrew Drucker. High-confidence predictions under adversarial uncertainty. *ACM Transactions on Computation Theory*, 5(3), August 2013. `doi:10.1145/2493252.2493257`.

14 Pavol Ďuriš. A note on the hierarchy of one-way data-independent multi-head finite automata, 2012. Technical report. URL: `https://eccc.weizmann.ac.il/report/2012/092/`.

15 Pavol Ďuriš, Rastislav Královič, and Dana Pardubská. Tight hierarchy of data-independent multi-head automata. *Journal of Computer and System Sciences*, 114:126–136, 2020. `doi:10.1016/j.jcss.2020.06.005`.

16 Lance Fortnow and Jack H. Lutz. Prediction and dimension. *Journal of Computer and System Sciences*, 70(4):570–589, 2005. Special Issue on COLT 2002. `doi:10.1016/j.jcss.2004.10.007`.

17 Xiaoyang Gu, Jack H. Lutz, Elvira Mayordomo, and Philippe Moser. Dimension spectra of random subfractals of self-similar fractals. *Annals of Pure and Applied Logic*, 165(11):1707–1726, 2014. `doi:10.1016/j.apal.2014.07.001`.

18 Felix Hausdorff. Dimension und äußeres Maß. *Mathematische Annalen*, 79:157–179, 1919. `doi:10.1007/BF01457179`.

19 John Hitchcock. Effective fractal dimension bibliography. `https://www.cs.uwyo.edu/~jhitchco/bib/dim`, accessed June 2026.

20 John M. Hitchcock. Fractal dimension and logarithmic loss unpredictability. *Theoretical Computer Science*, 304(1):431–441, 2003. `doi:10.1016/S0304-3975(03)00138-5`.

21 John M. Hitchcock and N. V. Vinodchandran. Dimension, entropy rates, and compression. *Journal of Computer and System Sciences*, 72(4):760–782, 2006. `doi:10.1016/j.jcss.2005.10.002`.

22 Markus Holzer. *Data-independent versus data-dependent computations on multi-head automata*. PhD thesis, Universität Tübingen, 1998.

23 Markus Holzer. Multi-head finite automata: data-independent versus data-dependent computations. *Theoretical Computer Science*, 286(1):97–116, 2002. `doi:10.1016/S0304-3975(01)00237-7`.

24 Markus Holzer, Martin Kutrib, and Andreas Malcher. Multi-head finite automata: Characterizations, concepts and open problems. In *Electronic Proceedings in Theoretical Computer Science (EPTCS)*, volume 1, pages 93–107, 2009. `doi:10.4204/EPTCS.1.9`.

25 Xiang Huang, Jack H. Lutz, Elvira Mayordomo, and Donald M. Stull. Asymptotic divergences and strong dichotomy. *IEEE Transactions on Information Theory*, 67(10):6296–6305, 2021. `doi:10.1109/TIT.2021.3085425`.

26 Alexander Kozachinskiy and Alexander Shen. Two characterizations of finite-state dimension. In Leszek Antoni Gąsieniec, Jesper Jansson, and Christos Levcopoulos, editors, *Fundamentals of Computation Theory*, pages 80–94, Cham, 2019. Springer International Publishing. `doi:10.1007/978-3-030-25027-0_6`.

27 Alexander Kozachinskiy and Alexander Shen. Automatic Kolmogorov complexity, normality, and finite-state dimension revisited. *Journal of Computer and System Sciences*, 118:75–107, 2021. `doi:10.1016/j.jcss.2020.12.003`.

28 Ming Li and Paul M. B. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer-Verlag, Berlin, fourth edition, 2019. `doi:10.1007/978-3-030-11298-1`.

29 María López-Valdés. Lempel–Ziv dimension for Lempel–Ziv compression. In Rastislav Královič and Paweł Urzyczyn, editors, *Mathematical Foundations of Computer Science 2006*, Lecture Notes in Computer Science, pages 693–703. Springer, 2006. `doi:10.1007/11821069_60`.

30 Jack H. Lutz. Dimension in complexity classes. *SIAM Journal on Computing*, 32(5):1236–1259, 2003. doi:10.1137/S0097539701417723.

31 Jack H. Lutz. The dimensions of individual strings and sequences. *Information and Computation*, 187(1):49–79, 2003. doi:10.1016/S0890-5401(03)00187-1.

32 Jack H. Lutz. A divergence formula for randomness and dimension. *Theoretical Computer Science*, 412:166–177, 2011. doi:10.1016/j.tcs.2010.09.005.

33 Jack H. Lutz and Elvira Mayordomo. Computing absolutely normal numbers in nearly linear time. *Information and Computation*, 281:104746, 2021. doi:10.1016/j.ic.2021.104746.

34 Neil Lutz. Multi-head finite-state compression. In *LATIN 2026: Theoretical Informatics*, Lecture Notes in Computer Science. Springer, to appear.

35 Elvira Mayordomo, Philippe Moser, and Sylvain Perifel. Polylog space compression, pushdown compression, and Lempel–Ziv are incomparable. *Theory of Computing Systems*, 48:731–766, 2011. doi:10.1007/s00224-010-9267-6.

36 Burkhard Monien. Two-way multihead automata over a one-letter alphabet. *RAIRO. Informatique théorique*, 14(1):67–82, 1980. doi:10.1051/ita/1980140100671.

37 Satyadev Nandakumar and Subin Pulari. Real numbers equally compressible in every base. *ACM Transactions on Computation Theory*, 17(3), June 2025. doi:10.1145/3729207.

38 Satyadev Nandakumar, Subin Pulari, and Akhil S. Finite-state relative dimension, dimensions of A. P. subsequences and a finite-state van Lambalgen's theorem. *Information and Computation*, 298:105156, 2024. doi:10.1016/j.ic.2024.105156.

39 Satyadev Nandakumar and Santhosh Kumar Vangapelli. Normality and finite-state dimension of Liouville numbers. *Theory of Computing Systems*, 58(3):392–402, 2016. doi:10.1007/s00224-014-9554-8.

40 Satyadev Nandakumar and Prateek Vishnoi. Randomness and Effective Dimension of Continued Fractions. In Javier Esparza and Daniel Král', editors, *45th International Symposium on Mathematical Foundations of Computer Science (MFCS 2020)*, volume 170 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 73:1–73:13, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.MFCS.2020.73.

41 Jared Nichols. Pushdown gamblers and pushdown dimension. Master's thesis, Iowa State University, Ames, IA, January 2004. URL: https://dr.lib.iastate.edu/handle/20.500.12876/97583.

42 Nicholas Pippenger and Michael J. Fischer. Relations among complexity measures. *Journal of the ACM*, 26(2):361–381, April 1979. doi:10.1145/322123.322138.

43 M. O. Rabin and D. Scott. Finite automata and their decision problems. *IBM Journal of Research and Development*, 3(2):114–125, 1959. doi:10.1147/rd.32.0114.

44 Arnold L. Rosenberg. On multi-head finite automata. *IBM Journal of Research and Development*, 10(5):388–394, 1966. doi:10.1147/rd.105.0388.

45 C.-P. Schnorr and H. Stimm. Endliche automaten und zufallsfolgen. *Acta Informatica*, 1(4):345–359, 1972. doi:10.1007/BF00289514.

46 D. Sullivan. Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups. *Acta Mathematica*, 153:259–277, 1984. doi:10.1007/BF02392379.

47 Claude Tricot. Two definitions of fractional dimension. *Mathematical Proceedings of the Cambridge Philosophical Society*, 91:57–74, 1982. doi:10.1017/S0305004100059119.

48 Andrew C. Yao and Ronald L. Rivest. $k+1$ heads are better than k. *Journal of the ACM*, 25(2):337–340, 1978. doi:10.1145/322063.322076.