Finite-State Dimension for Continued Fractions: Betting, Entropy and Normality

Satyadev Nandakumar
Department of Computer Science and Engineering, Indian Institute of Technology Kanpur, India

Subin Pulari
National Research University Higher School of Economics Moscow, Russia

Akhil S
Department of Computer Science and Engineering, Indian Institute of Technology Kanpur, India

Abstract

Finite-state dimension quantifies the asymptotic density of information in an infinite sequence as seen by finite automata, and can be viewed as a bounded-memory analogue of Hausdorff dimension. The theory is by now well developed for base-b expansions. In this paper we initiate the study of finite-state dimension in the setting of continued fractions. This setting brings several new difficulties. The natural reference measure is the Gauss measure, which is the canonical invariant probability measure for the Gauss transformation governing the continued fraction shift. Unlike in the base-b setting, however, this measure is not a product measure. In addition, the digit space is inherently infinite, so the symbolic setting is markedly less rigid than the usual finite-alphabet framework. Continued fraction analogues of effective dimension have received considerable attention in recent years, but a finite-state theory in this setting has so far been missing.

Finite-state dimension in the classical setting admits several equivalent formulations. We begin from the original finite-state s-gale viewpoint. In the continued fraction setting, however, the betting odds are no longer fixed in advance: unlike in the base-b case, the fair payoff for the next symbol is determined by conditional Gauss probabilities, and these vary with the current continued fraction cylinder. This makes the choice of a finite-state betting model genuinely nontrivial. We first examine a natural local finite-state betting model for truncated continued fraction digits, in which a gambler may use both the visible local context and an additional finite internal memory. We show that this model is too strong: there is a fixed finite-state gambler that wins at a positive exponential rate on every continued fraction normal point. Consequently, full dimension in this model does not characterize continued fraction normality, and the classical Schnorr–Stimm dichotomy fails.

We then isolate the source of this failure and introduce a restricted context-gambler model in which the gambler is allowed to use only the visible truncated context, with no additional hidden finite-state memory. For this restricted model, we obtain, in direct analogy with the base-b setting, an entropy-rate characterization of finite-state dimension in terms of conditional entropy rates, and from this derive an exact finite-state characterization of continued fraction normality: an irrational real has finite-state dimension 1 if and only if its continued fraction expansion is normal. Finally, we prove that the Schnorr–Stimm dichotomy does hold in this restricted setting: on a continued fraction normal point every such gambler either preserves constant capital or loses at an exponential rate, while on every non-normal point some such gambler wins at an exponential rate.

2012 ACM Subject Classification Theory of computation → Computability; Mathematics of computing → Information theory

Keywords and phrases Finite-state dimension, continued fractions, normality, finite-state gamblers, entropy rates

Digital Object Identifier 10.4230/LIPIcs.MFCS.2026.39

Funding *Subin Pulari*: supported by the Basic Research Program at HSE University and this paper was prepared in the framework of this program.

51st International Symposium on Mathematical Foundations of Computer Science (MFCS 2026).
Editors: Michal Koucký and Daniela Petrişan; Article No. 39; pp. 39:1–39:17
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Finite-state dimension is a quantitative notion of randomness measured using finite-state automata. Introduced by Dai, Lathrop, Lutz, and Mayordomo as a finite-state analogue of Hausdorff dimension [7], it measures the lower asymptotic density of information in an infinite sequence as detected by finite-state automata, and thus isolates exactly the amount of structure visible to bounded-memory computation. It admits several robust equivalent formulations, including ones via finite-state s-gales, block entropy rates, prediction, and finite-state compression [7, 5, 11]. From the gambling viewpoint, finite-state dimension is defined using finite-state s-gales. A finite-state gambler reads the sequence left to right, keeps only finitely many internal states, and in each state q chooses a betting distribution β_q on the next digit. In the base-b setting the reference measure is the uniform product measure, so after any prefix w the fair conditional probability of seeing the next symbol a is $\frac{\mu([wa])}{\mu([w])} = \frac{1}{b}$. Accordingly, if the gambler is in state q and bets according to β_q, the capital update at exponent s is

$$\frac{d(wa)}{d(w)} = \frac{\beta_q(a)}{\big(\mu([wa])/\mu([w])\big)^s} = \frac{\beta_q(a)}{(1/b)^s}.$$

The parameter s controls how favorable the game is to the gambler: larger s makes capital growth easier, while smaller s requires stronger detectable bias in the sequence. A gambler succeeds at exponent s if its capital is unbounded along the sequence, and the finite-state dimension is the infimum of the exponents $s \geq 0$ for which some finite-state gambler succeeds. In the classical base-b setting, this notion aligns perfectly with Borel normality. If $x = 0.x_1x_2x_3\cdots$ with $x_i \in \{0, 1, \dots, b-1\}$, then x is normal in base b precisely when every block u of length m appears with the expected limiting frequency

$$\lim_{n\to\infty} \frac{1}{n}\big|\{0 \leq i < n : x_{i+1}x_{i+2}\cdots x_{i+m} = u\}\big| = b^{-m}.$$

A real number is normal in base b if and only if its base-b expansion has finite-state dimension 1 [5, 21]. Moreover, the Schnorr–Stimm dichotomy theorem [21] gives a sharper dynamical picture: on a normal sequence every finite-state gambler either keeps its capital constant or loses at an exponential rate, while on every non-normal sequence some finite-state gambler wins at an exponential rate.

In this paper we initiate the study of finite-state randomness in the setting of continued fractions. Every irrational real $x \in (0, 1)$ has a unique continued fraction expansion

$$x = [0; a_1, a_2, a_3, \dots] = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}},$$

where the digits a_i are positive integers. Continued fractions are one of the most natural symbolic representations of real numbers in number theory and dynamics, and continued fraction normality is defined with respect to the Gauss measure [9, 13]. Concretely, x is continued fraction normal if for every finite block $c_1, \dots, c_m$,

$$\lim_{n\to\infty} \frac{1}{n}\big|\{0 \leq i < n : (a_{i+1}, \dots, a_{i+m}) = (c_1, \dots, c_m)\}\big| = \gamma(C_{c_1\cdots c_m}),$$

where $C_{c_1\cdots c_m}$ is the corresponding continued fraction cylinder. Here γ denotes the Gauss measure, given by $\gamma(A) = \frac{1}{\log 2}\int_A \frac{dx}{1+x}$. This is the natural reference measure for continued fractions because it is invariant under the Gauss transformation $T(x) = \frac{1}{x} - \lfloor 1/x \rfloor$, which acts as the left shift on continued fraction expansions. Several explicit constructions of continued fraction normal numbers are also known [1, 4]. It is therefore natural to ask for a finite-state theory that plays, for continued fractions, the role that finite-state dimension plays for base-b expansions.

Several aspects of the continued fraction setting are immediately different. First, the digit alphabet is countably infinite rather than finite, so unlike in the usual finite-alphabet setting, some additional modeling choice is needed in order to formulate a finite-state theory. In this paper we do this by truncating the digits at a threshold K and collapsing all larger digits into a single overflow symbol. Second, the relevant reference measure is the Gauss measure, the canonical invariant measure for the continued fraction system, and it is not a product measure.[1] In particular, unlike in the base-b case, the law of the next digit is not fixed in advance. In the base-b setting, if μ denotes the uniform product measure, then

$$\frac{\mu([wa])}{\mu([w])} = \frac{1}{b},$$

independently of the past w. In the continued fraction setting, by contrast, the fair payoff for betting on the next truncated symbol after a prefix w is governed by

$$\frac{\gamma(C_{wa})}{\gamma(C_w)},$$

and this conditional probability genuinely varies with the current cylinder. This moving-payoff phenomenon is the basic obstruction that makes the finite-state theory nontrivial.

Several recent works studied continued fraction analogues of effective randomness and effective dimension [16, 19, 18, 23]. This line of work shows that continued fractions support a rich and subtle algorithmic theory, but it does not address the finite-state regime. Our aim here is to seek a notion of dimension that captures exactly the randomness detectable by finite automata in continued fraction expansions, and in particular recovers continued fraction normality at dimension 1.

We take as our starting point the finite-state s-gale viewpoint [7]. In the base-b setting, the update may be written using the ratio $\mu([wa])/\mu([w])$, but since μ is a product measure this is just the fixed one-step law $1/b$. In the continued fraction setting, by contrast, the corresponding conditional probability depends genuinely on the past. This leads to the following natural first model: after reading a prefix w, the gambler reaches a state q and bets according to a distribution β_q, with capital update

$$d(wa) = d(w)\,\frac{\beta_q(a)}{\big(\gamma(C_{wa})/\gamma(C_w)\big)^s}.$$

We call this the *full-past* model. Although mathematically natural, it is not the bounded-memory notion relevant here, because the payoff rule itself depends on the entire prefix. Continued fraction normality is instead a finite-block frequency notion. In the base-b setting this distinction disappears because the one-step law is constant; in the continued fraction setting it does not, so the full-past model is tied to the whole history rather than to finite-state local behavior.[2]

The next natural step is to localize the changing payoff rule. We therefore pass to truncated continued fraction digits and let the fair payoff depend only on a bounded visible context, while still allowing the gambler an additional finite internal state. Write $\gamma_K(w)$

[1] See the preliminaries for an example.
[2] See Section 8 for further discussion on the full-past model.

for the Gauss measure of the truncated cylinder determined by a finite truncated word w. Concretely, after a truncated prefix w, let v denote its current visible suffix of some fixed length r, and let q be the gambler's internal state. The gambler then bets according to a distribution β_q on the next symbol, while the fair payoff is determined by the reference conditional probability of that next truncated symbol given the visible context v. The update rule therefore takes the form

$$d(wa) = d(w)\,\frac{\beta_q(a)}{\big(\gamma_K(va)/\gamma_K(v)\big)^s}.$$

where the denominator uses only the visible context, but the numerator may also exploit the hidden state q. One might expect this to be the appropriate continued fraction analogue of ordinary finite-state gamblers. We show that it is not.

In fact, this model is too strong: there is a fixed gambler of this type that wins at a positive exponential rate on every continued fraction normal point. If a gambler already grows like e^{cn} at exponent 1 for some $c > 0$, then the same strategy still has positive asymptotic growth for all $s < 1$ sufficiently close to 1. Consequently, any notion of dimension built from this class of gamblers fails to characterize continued fraction normality: even continued fraction normal reals can have dimension below 1 in that sense. The same counterexample also shows that the classical Schnorr–Stimm dichotomy fails in this setting.

The failure is not accidental. It stems from the non-product nature of the Gauss measure: the gambler's finite-state memory can retain information that is invisible to the payoff rule, and this mismatch creates systematic gain even on normal sequences. This issue is absent in the base-b setting, where the reference measure is product and the next-symbol payoff is independent of the past. We therefore introduce a *restricted context-gambler* model in which the gambler uses only the visible truncated context, with no additional finite-state memory, so that the information used in the betting rule and the fair-payoff rule are aligned.

For this restricted model, the classical finite-state picture re-emerges. First, in direct analogy with the base-b setting, we obtain an entropy-rate formulation of finite-state dimension in terms of conditional entropy rates. Second, we prove that an irrational real has finite-state dimension 1 if and only if its continued fraction expansion is normal. The converse direction uses, in a controlled way, the finite Markov chain induced by the truncated process together with a stationary-distribution argument. Finally, we establish a Schnorr–Stimm dichotomy for the restricted model: on a continued fraction normal point every such gambler either matches the reference conditional law and preserves constant capital, or else loses at an exponential rate; on every non-normal point some such gambler wins at an exponential rate.

The main contribution of the paper is the identification of the exact gambler model for which the classical finite-state theory extends correctly to the continued fraction setting. In this sense, the paper provides the first finite-state randomness theory for continued fractions that simultaneously respects the Gauss dynamics, admits an entropy-rate formulation, characterizes continued fraction normality by finite-state dimension 1, and satisfies the expected Schnorr–Stimm dichotomy.

2 Preliminaries

We write $\mathbb{N} = \{1, 2, 3, \dots\}$ and λ for the empty word. Every irrational real $x \in (0, 1)$ has a unique continued fraction expansion $x = [0; a_1(x), a_2(x), a_3(x), \dots]$. For a finite word $w = (w_1, \dots, w_m) \in \mathbb{N}^m$, we write $C_w = \{x \in (0,1) \setminus \mathbb{Q} : a_1(x) = w_1, \dots, a_m(x) = w_m\}$ for the corresponding continued fraction cylinder. We denote by μ the Lebesgue measure on $(0, 1)$, and by γ the Gauss measure [14, 9],

$$\gamma(A) = \frac{1}{\log 2}\int_A \frac{dx}{1+x}.$$

The Gauss transformation is $T(x) = \frac{1}{x} - \lfloor 1/x \rfloor$. It acts as the left shift on continued fraction expansions: if $x = [0; a_1, a_2, \ldots]$, then $T(x) = [0; a_2, a_3, \ldots]$. The measure γ is the canonical invariant probability measure for this transformation [9, 13]. A real $x \in (0,1) \setminus \mathbb{Q}$ is called *continued fraction normal* if for every finite word $w \in \mathbb{N}^m$,

$$\lim_{n\to\infty} \frac{1}{n} \big|\{0 \le i < n : T^i x \in C_w\}\big| = \gamma(C_w).$$

The Gauss measure is not a product measure. Indeed, if $C_{[1]}$ and $C_{[1,1]}$ denote the cylinders corresponding to the blocks 1 and 11, then $\gamma(C_{[1]}) = \log(4/3)/\log 2$ and $\gamma(C_{[1,1]}) = \log(10/9)/\log 2$, so $\gamma(C_{[1,1]}) \ne \gamma(C_{[1]})^2$.

To obtain a finite observable alphabet, we fix $K \ge 1$ and define $\Sigma_K = \{1, 2, \ldots, K, *\}$. The truncation map $\tau_K : \mathbb{N} \to \Sigma_K$ is given by

$$\tau_K(n) = \begin{cases} n, & 1 \le n \le K, \\ *, & n > K. \end{cases}$$

For $x = [0; a_1, a_2, \ldots]$, we write $y^{(K)}(x) = \tau_K(a_1)\tau_K(a_2)\tau_K(a_3)\cdots \in \Sigma_K^{\mathbb{N}}$ for the truncated digit sequence. If $u = u_1 \cdots u_m \in \Sigma_K^m$, we define the truncated cylinder

$$C_u^{(K)} = \{x \in (0,1) \setminus \mathbb{Q} : \tau_K(a_1(x)) = u_1, \ldots, \tau_K(a_m(x)) = u_m\},$$

and set $\gamma_K(u) = \gamma(C_u^{(K)})$.

▶ **Lemma 1.** *For every $K \ge 1$ and every finite word $u \in \Sigma_K^m$, one has $\gamma_K(u) > 0$.*

Fix $r \ge 0$. If $r = 0$, let $V_{K,0} = \{\lambda\}$, and if $r \ge 1$, let $V_{K,r} = \Sigma_K^r$. For a finite word $u \in \Sigma_K^*$, we write $\mathrm{suf}_r(u)$ for its suffix of length r when $r \ge 1$, and let $\mathrm{suf}_0(u) = \lambda$. For $v \in V_{K,r}$ and $a \in \Sigma_K$, define

$$p_{K,r}(a \mid v) = \frac{\gamma_K(va)}{\gamma_K(v)}.$$

Thus $p_{K,r}(\cdot \mid v)$ is the order-r truncated conditional law induced by the Gauss measure.

▶ **Lemma 2.** *For every fixed $K \ge 1$ and $r \ge 0$, every $v \in V_{K,r}$, and every $a \in \Sigma_K$, one has $p_{K,r}(a \mid v) > 0$, and also $\sum_{a \in \Sigma_K} p_{K,r}(a \mid v) = 1$. Moreover, if $M_{K,r} = \max\{p_{K,r}(a \mid v) : v \in V_{K,r},\ a \in \Sigma_K\}$, then $0 < M_{K,r} < 1$.*

The following elementary lemma shows that continued fraction normality is preserved when one passes to truncated digit sequences.

▶ **Lemma 3** (Truncation preserves normality)**.** *Let $x = [0; a_1, a_2, \ldots]$ be continued fraction normal. Fix $K \ge 1$ and a truncated word $u \in \Sigma_K^m$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \big|\{0 \le i < n : \tau_K(a_{i+1}) \cdots \tau_K(a_{i+m}) = u\}\big| = \gamma_K(u).$$

For a finite set A, we write $\mathcal{P}(A)$ for the simplex of all real-valued probability distributions on A. If $\rho \in \mathcal{P}(A)$, its Shannon entropy is $H(\rho) = -\sum_{a \in A} \rho(a) \log \rho(a)$, with the convention $0 \log 0 = 0$. If $\rho, \sigma \in \mathcal{P}(A)$ and $\sigma(a) > 0$ for every $a \in A$, we write $H(\rho, \sigma) = \sum_{a \in A} \rho(a) \log \frac{1}{\sigma(a)}$ for the cross-entropy of ρ relative to σ, and $\mathrm{D}(\rho \| \sigma) = \sum_{a \in A} \rho(a) \log \frac{\rho(a)}{\sigma(a)}$ for the Kullback–Leibler divergence. Thus $H(\rho, \sigma) = H(\rho) + \mathrm{D}(\rho \| \sigma)$. We shall use repeatedly the standard fact that $\mathrm{D}(\rho \| \sigma) \ge 0$, with equality if and only if $\rho = \sigma$ [6].

Finite-state dimension in the base-b setting

For comparison with the continued fraction setting, we briefly recall the classical notion of finite-state dimension in base b, following Dai, Lathrop, Lutz, and Mayordomo [7].[3] Fix an integer $b \geq 2$ and let $\Sigma_b = \{0, 1, \ldots, b-1\}$. A finite-state gambler over Σ_b is a quadruple $G = (Q, q_0, \delta, \beta)$, where Q is a finite set of states, $q_0 \in Q$ is the initial state, $\delta : Q \times \Sigma_b \to Q$ is the transition function, and $\beta : Q \to \mathcal{P}(\Sigma_b)$ is the betting function. Writing $\delta^*(q_0, w)$ for the state reached after reading a finite word $w \in \Sigma_b^*$, the associated s-gale is defined recursively by $d_G^{(s)}(\lambda) = 1$ and $d_G^{(s)}(wa) = d_G^{(s)}(w) b^s \beta(\delta^*(q_0, w))(a)$ for $w \in \Sigma_b^*$ and $a \in \Sigma_b$. Thus at each step the gambler redistributes its current capital according to the betting distribution prescribed by its current state, and the factor b^s determines the growth rate relevant to dimension. If $S \in \Sigma_b^{\mathbb{N}}$ is an infinite sequence, we say that G *succeeds at exponent s on S* if $\limsup_{n\to\infty} d_G^{(s)}(S\restriction n) = \infty$. The finite-state dimension of S is then

$$\dim_{\mathrm{FS}}(S) = \inf\{s \geq 0 : \text{some finite-state gambler succeeds at exponent } s \text{ on } S\}.$$

Intuitively, $\dim_{\mathrm{FS}}(S)$ measures how much randomness remains visible to finite automata: the smaller the value, the more regularity a finite-state gambler can exploit.

▶ **Theorem 4** ([21, 5]). *Let $b \geq 2$ and let $S \in \Sigma_b^{\mathbb{N}}$. Then S is normal in base b if and only if* $\dim_{\mathrm{FS}}(S) = 1$.

3 Natural betting models that do not capture continued fraction normality

We begin by considering two natural betting models that arise before one arrives at the restricted context model used later in the paper. The first is the most direct analogue of the usual gale update: after reading a truncated prefix $u \in \Sigma_K^*$, the finite-state gambler reaches some internal state q and then bets according to a probability distribution $\beta_q \in \mathcal{P}(\Sigma_K)$ on the next symbol. The fair payoff is given by the exact conditional Gauss probability $\gamma_K(ua)/\gamma_K(u)$, so the capital update takes the form

$$d(ua) = d(u)\,\frac{\beta_q(a)}{(\gamma_K(ua)/\gamma_K(u))^s}.$$

This model is mathematically natural, but it is not the finite-state notion sought here, since the fair-payoff rule itself depends on the entire prefix u and is therefore tied to full-past prediction rather than bounded-memory behaviour. In the base-b setting this issue is invisible, because the corresponding ratio is constantly $1/b$, independent of the past; in the continued fraction setting it is not. This indicates that some restriction on the past is necessary. The next natural candidate is therefore to replace the whole prefix by a bounded visible context, while still allowing the gambler a finite internal state. As we shall see, however, even this local finite-state model is too strong.

Fix $K \geq 1$ and $r \geq 0$. Write $\Sigma = \Sigma_K$, $V = V_{K,r}$, and $p(a \mid v) = p_{K,r}(a \mid v)$. An *unrestricted local finite-state gambler* at parameters (K, r) is a quadruple $G = (Q, q_0, \delta, \beta)$, where

[3] Dai et al. define finite-state dimension for a sequence via the infimum of the exponents s for which some finite-state s-gale succeeds on that sequence. The formulation used here is equivalent for individual sequences and is adopted only for convenience of presentation.

- Q is a nonempty finite set of internal states,
- $q_0 \in Q$ is the initial state,
- $\delta : Q \times \Sigma \to Q$ is the transition function,
- $\beta : Q \to \mathcal{P}(\Sigma)$ is the betting function.

Thus the gambler itself is an ordinary finite-state gambler in the sense of Dai et al. [7], while the local feature of the model lies in the fair-payoff rule, which is indexed by the visible suffix context. For a word $u \in \Sigma^*$, let $q_G(u)$ denote the state reached by G after reading u. For $s \geq 0$, the associated local s-capital process is defined by $d_G^{(s)}(u) = 1$ if $|u| = r$, and

$$d_G^{(s)}(ua) = d_G^{(s)}(u)\,\frac{\beta(q_G(u))(a)}{p(a \mid \mathrm{suf}_r(u))^s} \qquad (|u| \geq r,\ a \in \Sigma).$$

If $x \in (0,1) \setminus \mathbb{Q}$ and $y = y^{(K)}(x) = y_1 y_2 y_3 \cdots$, we say that G *succeeds at exponent s on x* if $\limsup_{n\to\infty} d_G^{(s)}(y_1 \cdots y_n) = \infty$.

Thus success means that the capital process is unbounded along the truncated digit sequence of x. This is exactly the model that looks natural at first sight, but it allows the gambler to use more information than the denominator uses. We now show that this unrestricted local model admits a fixed finite-state gambler that wins exponentially on *every* continued fraction normal point. We give this example in the case when $(K, r) = (1, 0)$.

For $K = 1$, we have $\Sigma_1 = \{1, *\}$. Since $r = 0$, the visible context set is $V_{1,0} = \{\lambda\}$, and the local reference probability is just the one-symbol marginal: $p_{1,0}(a \mid \lambda) = \gamma_1(a)$ for $a \in \Sigma_1$.

▶ **Theorem 5.** *There exists a fixed unrestricted local finite-state gambler G at parameters $(K, r) = (1, 0)$ such that for every continued fraction normal real number x,*

$$\lim_{n\to\infty} \frac{1}{n} \log d_G^{(1)}\big(y_1^{(1)}(x) \cdots y_n^{(1)}(x)\big) = I_1 > 0 \qquad \textit{where} \qquad I_1 = \sum_{b,a\in\Sigma_1} \gamma_1(ba) \log \frac{\gamma_1(ba)}{\gamma_1(b)\gamma_1(a)}.$$

In particular, the gambler G succeeds on every continued fraction normal point at some exponent $s_0 < 1$.

Proof sketch. Take a 3-state gambler G with states $\{\perp, 1, *\}$, where after the first step the state simply remembers the previous truncated digit. From state $b \in \Sigma_1$, let the gambler bet according to the conditional pair law $\beta(b)(a) = \gamma_1(ba)/\gamma_1(b)$, while at the initial state $\perp$ it bets according to the marginal γ_1. The key point is that for the unrestricted local model with $r = 0$, the payoff denominator still uses only the one-symbol marginal law $\gamma_1(a)$. Thus, after the first step, the capital multiplier at symbol y_i is

$$\frac{\gamma_1(y_{i-1} y_i)}{\gamma_1(y_{i-1})\gamma_1(y_i)^s}$$

so truncation normality gives $\frac{1}{n} \log d_G^{(s)}(y_1 \cdots y_n) \to L_1(s)$, where

$$L_1(s) = \sum_{b,a\in\Sigma_1} \gamma_1(ba) \log \frac{\gamma_1(ba)}{\gamma_1(b)\gamma_1(a)^s}.$$

At $s = 1$, this becomes $L_1(1) = \mathrm{D}\left(\gamma_1^{(2)} \,\|\, \gamma_1^{(1)} \otimes \gamma_1^{(1)}\right)$. Since $\gamma_1(C_{[1,1]}) \neq \gamma_1(C_{[1]})^2$, the pair law is not a product measure, so $L_1(1) > 0$. Hence G wins at a positive exponential rate on every continued fraction normal point. Finally, $L_1(s) = L_1(1) - (1 - s)H_1$, where $H_1 = -\sum_{a\in\Sigma_1} \gamma_1(a) \log \gamma_1(a) > 0$. Therefore $L_1(s) > 0$ for some $s < 1$ sufficiently close to 1, so the same gambler already succeeds at an exponent strictly below 1. ◀

4 Context gamblers

The counterexample from the previous section identifies the source of the problem with the unrestricted local model. There, the denominator used only the visible local context, while the gambler was allowed to retain additional hidden information in its internal state. Even at $(K,r)=(1,0)$, this mismatch lets the gambler use the previous truncated digit while the payoff rule sees only the one-symbol marginal law, and this creates a positive gain on continued fraction normal points. We therefore now restrict the gambler to use only the visible context and no additional hidden finite-state memory.

In this section we introduce the restricted model that will be used throughout the rest of the paper. We define context gamblers and their capital processes, record the exact gale identity they satisfy, and then define the associated notion of continued fraction dimension.

4.1 Definition

Fix (K,r) and abbreviate $\Sigma=\Sigma_K$, $V=V_{K,r}$, and $p(a \mid v)=p_{K,r}(a \mid v)$. We first isolate the betting object itself.

▶ **Definition 6.** *A* (K,r)-context gambler *is a function $\beta : V \to \mathcal{P}(\Sigma)$, that is, for each $v \in V$, $\beta(v)(\cdot)$ is a probability distribution on Σ.*

Intuitively, a (K,r)-context gambler is a finite-state gambler whose state information is exactly the currently visible truncated context: unlike in the unrestricted local model, its behaviour is determined entirely by that context. Given such a gambler, the corresponding capital process is defined by using exactly the same visible context in both the betting rule and the fair-payoff rule.

▶ **Definition 7.** *Fix $s \geq 0$. Let β be a (K,r)-context gambler. Its s-capital process $d_\beta^{(s)}$ is defined on words of length at least r by setting $d_\beta^{(s)}(u)=1$ whenever $|u|=r$, and*

$$d_\beta^{(s)}(ua)=d_\beta^{(s)}(u)\,\frac{\beta(\mathrm{suf}_r(u))(a)}{p(a \mid \mathrm{suf}_r(u))^s} \qquad (|u| \geq r,\ \ a \in \Sigma).$$

Because the numerator and denominator are now indexed by the same context, the capital process satisfies the following exact gale law.

▶ **Proposition 8** (Exact gale law)**.** *For every finite word $u \in \Sigma^*$ with $|u| \geq r$, $d_\beta^{(s)}(u)=\sum_{a\in\Sigma} p(a \mid \mathrm{suf}_r(u))^s\, d_\beta^{(s)}(ua)$.*

Proof.

$$\begin{aligned}\sum_{a\in\Sigma} p(a \mid \mathrm{suf}_r(u))^s\, d_\beta^{(s)}(ua) &= \sum_{a\in\Sigma} p(a \mid \mathrm{suf}_r(u))^s\, d_\beta^{(s)}(u)\,\frac{\beta(\mathrm{suf}_r(u))(a)}{p(a \mid \mathrm{suf}_r(u))^s}\\ &= d_\beta^{(s)}(u)\sum_{a\in\Sigma}\beta(\mathrm{suf}_r(u))(a)=d_\beta^{(s)}(u).\end{aligned}$$ ◀

Let $x \in (0,1)\setminus\mathbb{Q}$, let $y=y^{(K)}(x)=y_1y_2\cdots$, and let β be a (K,r)-context gambler. We say that β *succeeds at exponent s on x* if

$$\limsup_{n\to\infty} d_\beta^{(s)}(y_1\cdots y_n)=\infty.$$

This leads, for each fixed pair (K,r), to the quantity

$$\dim^{\mathrm{C}}_{K,r}(x)=\inf\{s \geq 0 : \text{some } (K,r)\text{-context gambler succeeds at exponent } s \text{ on } x\},$$

which measures the least exponent at which a gambler using truncation level K and context length r can succeed on x.

▶ **Definition 9.** *The* continued fraction finite-state dimension *of x is defined as*

$$\dim^{\mathrm{C}}_{\mathrm{CF}}(x) = \inf_{K\ge 1}\ \inf_{r\ge 0}\ \dim^{\mathrm{C}}_{K,r}(x).$$

5 Entropy-rate formulation of finite-state dimension

We now turn to the entropy formulation of the theory. Just as in the classical base-b setting [5], we obtain an entropy-rate characterization of the continued fraction finite-state dimension defined in the previous section. For each fixed truncation level K and context length r, we first define empirical conditional entropy quantities from the truncated expansion of x. These quantities measure, in a block-frequency sense, the uncertainty of the next truncated digit given the current visible context. We then show that the resulting entropy-rate quantity is exactly equivalent to the gambler-based notion.

Fix (K,r) and abbreviate again $\Sigma = \Sigma_K$, $V = V_{K,r}$, and $p(a \mid v) = p_{K,r}(a \mid v)$. Let $x \in (0,1)\backslash\mathbb{Q}$ and let $y = y^{(K)}(x) = y_1y_2\cdots$. For each $j \in \mathbb{N}$, define $v_j = \mathrm{suf}_r(y_1\cdots y_{r+j}) \in V$ and $a_j = y_{r+j+1} \in \Sigma$. Thus v_j is the current visible r-context immediately before the symbol a_j. For $n \ge 1$, define

$$N_n(v,a) = \big|\{0 \le j < n : (v_j,a_j) = (v,a)\}\big|, \quad P_n(v,a) = \frac{N_n(v,a)}{n}, \quad \alpha_n(v) = \sum_{a\in\Sigma} P_n(v,a).$$

Whenever $\alpha_n(v) > 0$, define the empirical conditional distribution by $P_n(a \mid v) = \frac{P_n(v,a)}{\alpha_n(v)}$. If $\alpha_n(v) = 0$, we define $P_n(\cdot \mid v)$ arbitrarily; such contexts carry zero weight in all the formulas below. Using the information-theoretic notation from the preliminaries, we now define

$$h^{K,r}_n(x) = \sum_{v\in V} \alpha_n(v)\, H\big(P_n(\cdot \mid v)\big) = \sum_{v\in V}\sum_{a\in\Sigma} P_n(v,a)\log\frac{1}{P_n(a \mid v)},$$

and

$$j^{K,r}_n(x) = \sum_{v\in V} \alpha_n(v)\, H\big(P_n(\cdot \mid v), p(\cdot \mid v)\big) = \sum_{v\in V}\sum_{a\in\Sigma} P_n(v,a)\log\frac{1}{p(a \mid v)}.$$

Thus $h^{K,r}_n(x)$ is the empirical conditional entropy of the next symbol given the current context, while $j^{K,r}_n(x)$ is the corresponding empirical cross-entropy relative to the reference probability $p_{K,r}$. We then define [4]

$$E_{K,r}(x) = \liminf_{n\to\infty} \frac{h^{K,r}_n(x)}{j^{K,r}_n(x)}.$$

▶ **Definition 10.** *The* continued fraction entropy dimension *of x is*

$$\dim^{\mathrm{E}}_{\mathrm{CF}}(x) = \inf_{K\ge 1}\ \inf_{r\ge 0}\ E_{K,r}(x).$$

The next lemma records the basic inequalities satisfied by these quantities and identifies the gap between $h^{K,r}_n(x)$ and $j^{K,r}_n(x)$ as an average Kullback–Leibler divergence.

▶ **Lemma 11.** *For every x, K, r, and $n \ge 1$, one has $0 \le h^{K,r}_n(x) \le j^{K,r}_n(x)$ and $j^{K,r}_n(x) \ge \log\frac{1}{M_{K,r}} > 0$. More precisely, $j^{K,r}_n(x) - h^{K,r}_n(x) = \sum_{v\in V} \alpha_n(v)\, \mathrm{D}\big(P_n(\cdot \mid v)\,\|\, p(\cdot \mid v)\big)$.*

[4] This is the continued fractions analogue of the usual normalized block entropy rate. In the base-b case the reference law is uniform, so the cross-entropy denominator is constant: for L-blocks it is $L\log b$. Thus the ratio $h^{K,r}_n/j^{K,r}_n$ reduces exactly to the standard block entropy rate in [5].

5.1 Equivalence of gamblers and entropy

We now prove that the entropy-rate formulation introduced above is exactly equivalent to the gambler-based notion of continued fraction finite-state dimension. We begin with the direction from entropy to gamblers.

▶ **Theorem 12.** *For every irrational x, every $K \geq 1$, and every $r \geq 0$, $\dim^{\mathrm{C}}_{K,r}(x) \leq E_{K,r}(x)$.*

Proof sketch. Fix $s > E_{K,r}(x)$. Choose a subsequence on which $h_n^{K,r}(x)/j_n^{K,r}(x)$ has limit $< s$, and pass further so that the empirical pair laws P_n converge to some P on $V \times \Sigma$. Continuity of the finite entropy expressions gives $s\, j(P) + \sum_{v,a} P(v,a) \log P(a \mid v) > 0$. We now build the gambler from this limiting law. In context v, it bets according to a small full-support perturbation of $P(\cdot \mid v)$. This removes zero-probability issues while preserving, by continuity, the above strict positivity for all large P_n along the subsequence. The log-capital identity then gives

$$\log d_\beta^{(s)}(y_1 \cdots y_{r+n}) = n \left(s\, j_n^{K,r}(x) + \sum_{v,a} P_n(v,a) \log \beta(v)(a) \right).$$

The bracketed term is eventually positive along the subsequence, so the capital grows exponentially there. Hence β succeeds at exponent s, giving $\dim^{\mathrm{C}}_{K,r}(x) \leq s$. Since this holds for every $s > E_{K,r}(x)$, the claim follows. ◀

▶ **Theorem 13.** *For every irrational x, every $K \geq 1$, and every $r \geq 0$, $E_{K,r}(x) \leq \dim^{\mathrm{C}}_{K,r}(x)$.*

Proof sketch. Fix $s > \dim^{\mathrm{C}}_{K,r}(x)$, and let β be a (K,r)-context gambler that succeeds at exponent s on x. For infinitely many n, its capital satisfies $d_\beta^{(s)}(y_1 \cdots y_{r+n}) > 1$. Writing the logarithm of the capital in terms of the empirical transition law P_n gives the inequality

$$s\, j_n^{K,r}(x) + \sum_{v,a} P_n(v,a) \log \beta(v)(a) > 0$$

for infinitely many n. The second term is controlled by the empirical conditional entropy: applying the nonnegativity of KL divergence separately in each context and then averaging with weights $\alpha_n(v)$ gives $\sum_{v,a} P_n(v,a) \log \beta(v)(a) \leq -h_n^{K,r}(x)$. Hence $h_n^{K,r}(x) < s\, j_n^{K,r}(x)$ along infinitely many n, and therefore $E_{K,r}(x) \leq s$. Since $s > \dim^{\mathrm{C}}_{K,r}(x)$ was arbitrary, $E_{K,r}(x) \leq \dim^{\mathrm{C}}_{K,r}(x)$. ◀

Together, the two theorems give the exact equivalence at each fixed pair of parameters.

▶ **Corollary 14.** *For every irrational x, every $K \geq 1$, and every $r \geq 0$, $\dim^{\mathrm{C}}_{K,r}(x) = E_{K,r}(x)$.*

Taking the infimum over all truncation levels and context lengths now yields the entropy-rate formulation of continued fraction finite-state dimension.

▶ **Corollary 15.** *For every irrational x, $\dim^{\mathrm{C}}_{\mathrm{CF}}(x) = \dim^{\mathrm{E}}_{\mathrm{CF}}(x)$.*

6 A dimension characterization of continued fraction normality

We now prove that full continued fraction finite-state dimension is exactly the same as continued fraction normality. First we show that continued fraction normality forces entropy dimension to be equal to 1. For the converse, starting from a failure of normality, we pass to a limiting empirical law on contexts and next symbols. The consistency relations between block frequencies forces its context marginal to be the stationary distribution for the finite Markov chain induced by the conditional probabilities and yields a positive entropy defect.

▶ **Theorem 16.** *If x is continued fraction normal, then* $\dim^{\mathrm{E}}_{\mathrm{CF}}(x) = 1$.

Proof. Fix arbitrary $K \geq 1$ and $r \geq 0$, and let $y = y^{(K)}(x)$. For each $v \in V_{K,r}$ and each $a \in \Sigma_K$, the event $(v_j, a_j) = (v, a)$ is exactly the event that the truncated word va occurs in y starting at position $j+1$. Therefore

$$P_n(v,a) = \frac{1}{n}\big|\{1 \leq i \leq n : y_i \cdots y_{i+r} = va\}\big|.$$

By Lemma 3, we have $P_n(v,a) \to \gamma_K(va)$ as $n \to \infty$ for every $v \in V_{K,r}$ and every $a \in \Sigma_K$. Consequently, $\alpha_n(v) = \sum_a P_n(v,a) \to \gamma_K(v)$ and $P_n(a \mid v) \to \gamma_K(va)/\gamma_K(v) = p_{K,r}(a \mid v)$ for every $v \in V_{K,r}$. By continuity of entropy and cross-entropy on the finite simplex, $h_n^{K,r}(x) \to h^*$ and $j_n^{K,r}(x) \to j^*$, where

$$h^* = \sum_{v \in V} \gamma_K(v) \sum_{a \in \Sigma} p(a \mid v) \log \frac{1}{p(a \mid v)}$$

and

$$j^* = \sum_{v \in V} \gamma_K(v) \sum_{a \in \Sigma} p(a \mid v) \log \frac{1}{p(a \mid v)}.$$

Thus $h^* = j^*$. By Lemma 11, we have $j_n^{K,r}(x) \geq \log(1/M_{K,r}) > 0$ for all n, so in particular $j^* > 0$. Therefore

$$E_{K,r}(x) = \lim_{n \to \infty} \frac{h_n^{K,r}(x)}{j_n^{K,r}(x)} = 1.$$

Since K and r were arbitrary, it follows that $\dim^{\mathrm{E}}_{\mathrm{CF}}(x) = \inf_{K \geq 1} \inf_{r \geq 0} E_{K,r}(x) = 1$. ◀

▶ **Theorem 17.** *If* $\dim^{\mathrm{E}}_{\mathrm{CF}}(x) = 1$, *then x is continued fraction normal.*

Proof. We prove the contrapositive. Assume that x is *not* continued fraction normal. Then there exists a finite ordinary word $b = b_1 \cdots b_t \in \mathbb{N}^t$ such that

$$\frac{1}{N}\big|\{0 \leq i < N : T^i x \in C_b\}\big|$$

does not converge to $\gamma(C_b)$ as $N \to \infty$. Choose $K \geq \max\{b_1, \ldots, b_t\}$. Then b may also be regarded as a word in Σ_K^t with no occurrence of $*$. Let $y = y^{(K)}(x) = y_1 y_2 \cdots$. Because every digit in b is at most K, an occurrence of b in the continued fraction digit expansion of x is exactly the same thing as an occurrence of the truncated word b in y. Therefore

$$\frac{1}{N}\big|\{1 \leq i \leq N : y_i \cdots y_{i+t-1} = b\}\big|$$

does not converge to $\gamma_K(b) = \gamma(C_b)$. Set $r = t-1$. Then for $v \in V_{K,r}$ and $a \in \Sigma_K$,

$$P_n(v,a) = \frac{1}{n}\big|\{1 \leq i \leq n : y_i \cdots y_{i+r} = va\}\big|.$$

In particular, writing $v^* = b_1 \cdots b_{t-1}$ and $a^* = b_t$, we have

$$P_n(v^*,a^*) = \frac{1}{n}\big|\{1 \leq i \leq n : y_i \cdots y_{i+t-1} = b\}\big|,$$

which does not converge to $\gamma_K(b)$. Therefore the family $\big(P_n(v,a)\big)_{v \in V_{K,r},\, a \in \Sigma_K}$ does not converge to $\big(\gamma_K(va)\big)_{v \in V_{K,r},\, a \in \Sigma_K}$. Since the simplex of probability distributions on the finite set $V_{K,r} \times \Sigma_K$ is compact, there exists a subsequence (n_m) and a probability distribution

$\nu \in \mathcal{P}(V_{K,r} \times \Sigma_K)$ such that $P_{n_m}(v,a) \to \nu(v,a)$ for every $v \in V_{K,r}$ and every $a \in \Sigma_K$, and such that $\nu(v^*,a^*) \neq \gamma_K(b)$. Define $\alpha(v) = \sum_{a\in\Sigma_K} \nu(v,a)$. If $\alpha(v) > 0$, define $\nu(a \mid v) = \nu(v,a)/\alpha(v)$; if $\alpha(v) = 0$, define $\nu(\cdot \mid v)$ arbitrarily. We now pass to the limiting entropy quantities. Since

$$h^{K,r}_{n_m}(x) = -\sum_{v,a} P_{n_m}(v,a) \log P_{n_m}(v,a) + \sum_{v} \alpha_{n_m}(v) \log \alpha_{n_m}(v),$$

continuity on the finite simplex gives

$$h^{K,r}_{n_m}(x) \longrightarrow h(\nu) := -\sum_{v,a} \nu(v,a) \log \nu(v,a) + \sum_{v} \alpha(v) \log \alpha(v) = \sum_{v,a} \nu(v,a) \log \frac{1}{\nu(a \mid v)}.$$

Also, $j^{K,r}_{n_m}(x) \to j(\nu) := \sum_{v,a} \nu(v,a) \log \frac{1}{p_{K,r}(a|v)}$, since $j^{K,r}_{n}(x)$ is linear in the empirical distribution. Moreover, Lemma 11 gives $j^{K,r}_{n}(x) \geq \log \frac{1}{M_{K,r}} > 0$ for all n, so in particular $j(\nu) \geq \log \frac{1}{M_{K,r}} > 0$. By Lemma 11,

$$j(\nu) - h(\nu) = \sum_{v \in V_{K,r}} \alpha(v)\; \mathrm{D}\big(\nu(\cdot \mid v) \,\|\, p_{K,r}(\cdot \mid v)\big).$$

We claim that $j(\nu) - h(\nu) > 0$. Suppose for contradiction that $j(\nu) - h(\nu) = 0$. Then every term in the sum is zero, so for every v with $\alpha(v) > 0$, we have $\nu(\cdot \mid v) = p_{K,r}(\cdot \mid v)$. Hence

$$\nu(v,a) = \alpha(v)\, p_{K,r}(a \mid v) \qquad \text{for every } v \in V_{K,r},\ a \in \Sigma_K.$$

We now split into two cases.

Case 1: $r = 0$. Then $V_{K,0} = \{\lambda\}$ and $\alpha(\lambda) = 1$. Therefore $\nu(\lambda,a) = p_{K,0}(a \mid \lambda) = \gamma_K(a)$ for all $a \in \Sigma_K$, so $\nu = \gamma_K$ on $V_{K,0} \times \Sigma_K$, contradicting the choice of ν.

Case 2: $r \geq 1$. Fix $v = v_1 \cdots v_r \in V_{K,r}$. Then $\sum_{a\in\Sigma_K} P_n(v,a) = \frac{1}{n}\big|\{1 \leq i \leq n : y_i \cdots y_{i+r-1} = v\}\big|$, while

$$\sum_{c \in \Sigma_K} P_n(cv_1 \cdots v_{r-1}, v_r) = \frac{1}{n}\big|\{2 \leq i \leq n+1 : y_i \cdots y_{i+r-1} = v\}\big|.$$

These two counts differ by at most 1, so

$$\left| \sum_{a \in \Sigma_K} P_n(v,a) - \sum_{c \in \Sigma_K} P_n(cv_1 \cdots v_{r-1}, v_r) \right| \leq \frac{1}{n}.$$

Passing to the subsequence limit gives

$$\alpha(v) = \sum_{c \in \Sigma_K} \nu(cv_1 \cdots v_{r-1}, v_r).$$

Using $\nu(v,a) = \alpha(v) p_{K,r}(a \mid v)$, we obtain

$$\alpha(w) = \sum_{v \in V_{K,r}} \sum_{\substack{a \in \Sigma_K \\ \mathrm{suf}_r(va) = w}} \alpha(v)\, p_{K,r}(a \mid v) \qquad (w \in V_{K,r}).$$

Thus α is a stationary distribution for the finite Markov chain on $V_{K,r}$ whose transition rule is: from v, output a with probability $p_{K,r}(a \mid v)$ and move to $\mathrm{suf}_r(va)$.

This chain is irreducible, because every one-step probability $p_{K,r}(a \mid v)$ is positive by Lemma 2, and from any state v one can reach any state $w = w_1 \cdots w_r$ by reading the symbols $w_1, \ldots, w_r$. Hence the chain has a unique stationary distribution.

On the other hand, the distribution $\pi(v) = \gamma_K(v)$ for $v \in V_{K,r}$ is stationary for the same chain. Indeed,

$$\sum_{v \in V_{K,r}} \sum_{\substack{a \in \Sigma_K \\ \mathrm{suf}_r(va) = w}} \pi(v)\, p_{K,r}(a \mid v) = \sum_{v \in V_{K,r}} \sum_{\substack{a \in \Sigma_K \\ \mathrm{suf}_r(va) = w}} \gamma_K(v) \frac{\gamma_K(va)}{\gamma_K(v)}$$
$$= \sum_{\substack{z \in \Sigma_K^{r+1} \\ \mathrm{suf}_r(z) = w}} \gamma_K(z).$$

Because the truncated process under γ is stationary, the probability that coordinates $2, \ldots, r+1$ equal w is the same as the probability that coordinates $1, \ldots, r$ equal w, namely $\gamma_K(w)$. The probability of the event "coordinates $2, \ldots, r+1$ equal w" is exactly the last displayed sum. Therefore

$$\sum_{v \in V_{K,r}} \sum_{\substack{a \in \Sigma_K \\ \mathrm{suf}_r(va) = w}} \pi(v)\, p_{K,r}(a \mid v) = \gamma_K(w) = \pi(w).$$

So π is stationary.

By uniqueness of stationary distributions, $\alpha = \pi$, that is, $\alpha(v) = \gamma_K(v)$ for every $v \in V_{K,r}$. Hence $\nu(v, a) = \alpha(v) p_{K,r}(a \mid v) = \gamma_K(v) \frac{\gamma_K(va)}{\gamma_K(v)} = \gamma_K(va)$ for every $v \in V_{K,r}$ and every $a \in \Sigma_K$. This contradicts the choice of ν.

In both cases we get a contradiction. Therefore $j(\nu) - h(\nu) > 0$. Since $j(\nu) > 0$, it follows that $h(\nu)/j(\nu) < 1$. But along the subsequence (n_m), we have

$$\frac{h_{n_m}^{K,r}(x)}{j_{n_m}^{K,r}(x)} \longrightarrow \frac{h(\nu)}{j(\nu)} < 1.$$

Therefore

$$E_{K,r}(x) = \liminf_{n \to \infty} \frac{h_n^{K,r}(x)}{j_n^{K,r}(x)} < 1.$$

Hence $\dim_{\mathrm{CF}}^{\mathrm{E}}(x) \leq E_{K,r}(x) < 1$.

This proves the contrapositive: if x is not continued fraction normal, then $\dim_{\mathrm{CF}}^{\mathrm{E}}(x) < 1$. Therefore $\dim_{\mathrm{CF}}^{\mathrm{E}}(x) = 1$ implies that x is continued fraction normal. ◀

Together with Corollary 15, the above results give the desired characterization.

▶ **Corollary 18.** *For every irrational $x \in (0, 1) \setminus \mathbb{Q}$,*

$$x \text{ is continued fraction normal} \iff \dim_{\mathrm{CF}}^{\mathrm{C}}(x) = 1 \iff \dim_{\mathrm{CF}}^{\mathrm{E}}(x) = 1.$$

At the other extreme, the dimension also identifies simple deterministic continued fraction expansions as having dimension zero. Recall that quadratic irrationals are exactly those real numbers whose continued fraction expansions are eventually periodic [14]. Let $x \in (0, 1) \setminus \mathbb{Q}$ be a quadratic irrational. Its continued fraction digits are eventually periodic, and hence bounded. Choose K large enough to contain all digits appearing in the expansion, and choose r large enough that every length-r context occurring after the initial prefix uniquely determines the next digit. For any $s > 0$, define a (K, r)-context gambler that bets probability 1 on this uniquely determined next digit on the contexts occurring along the eventual period, and define it arbitrarily elsewhere. Along the expansion of x, after the initial context, every bet is correct. Hence the capital grows exponentially for every $s > 0$. Since dimensions are nonnegative, it follows that

$$\dim_{\mathrm{CF}}^{\mathrm{C}}(x) = \dim_{\mathrm{CF}}^{\mathrm{E}}(x) = 0.$$

Thus all quadratic irrationals have continued fraction finite-state dimension zero.

7 A Schnorr–Stimm dichotomy for continued fractions

The previous sections introduced the restricted context-gambler model, established its entropy-rate formulation, and showed that continued fraction normality is equivalent to continued fraction finite-state dimension 1. In the classical base-b setting, a foundational theorem of Schnorr and Stimm [21] gives a sharp dynamical picture of finite-state randomness: on every normal sequence a finite-state gambler either preserves constant capital or loses at an exponential rate, while on every non-normal sequence some finite-state gambler wins at an exponential rate. The goal of this section is to prove the corresponding dichotomy for context gamblers in the continued fraction setting. A quantitative refinement of the classical Schnorr–Stimm picture was given by Huang, Lutz, Mayordomo, and Stull, who used asymptotic KL divergences and gambler risk to measure the exponential rates of winning and losing for finite-state gamblers over product reference measures [12].

We begin by recording the exact exponential growth rate of the capital process at exponent 1. Fix $K \geq 1$, $r \geq 0$, and a (K, r)-context gambler $\beta : V_{K,r} \to \mathcal{P}(\Sigma_K)$. Let $y = y^{(K)}(x) = y_1 y_2 y_3 \cdots$ be the truncated digit sequence of x. For $n \geq 1$, define $v_i = \mathrm{suf}_r(y_1 \cdots y_{r+i}) \in V_{K,r}$ and $a_i = y_{r+i+1} \in \Sigma_K$ for $0 \leq i < n$. Thus v_i is the visible context before the $(i+1)$st step after the initial r symbols, and a_i is the next symbol read at that step. Iterating the defining recursion for the capital process gives, for every $n \geq 1$,

$$\log d_\beta^{(1)}(y_1 \cdots y_{r+n}) = \sum_{i=0}^{n-1} \log \frac{\beta(v_i)(a_i)}{p_{K,r}(a_i \mid v_i)}. \tag{1}$$

For $v \in V_{K,r}$ and $a \in \Sigma_K$, let $P_n(v, a) = \frac{1}{n}\big|\{0 \leq i < n : (v_i, a_i) = (v, a)\}\big|$ be the empirical frequency of the pair (v, a). Grouping the sum in (1) by the pair (v, a) yields

$$\frac{1}{n} \log d_\beta^{(1)}(y_1 \cdots y_{r+n}) = \sum_{v \in V_{K,r}} \sum_{a \in \Sigma_K} P_n(v, a) \log \frac{\beta(v)(a)}{p_{K,r}(a \mid v)}. \tag{2}$$

We first analyze what happens on continued fraction normal points. The conclusion is that a context gambler either matches the reference distribution exactly and keeps its capital constant, or else it loses at an exponential rate.

▶ **Theorem 19.** *Let x be continued fraction normal, and fix $K \geq 1$, $r \geq 0$, and a (K, r)-context gambler β. Then*

$$\lim_{n \to \infty} \frac{1}{n} \log d_\beta^{(1)}(y_1 \cdots y_{r+n}) = -\sum_{v \in V_{K,r}} \gamma_K(v)\, \mathrm{D}\big(p_{K,r}(\cdot \mid v) \,\|\, \beta(v)\big).$$

In particular, $d_\beta^{(1)}(y_1 \cdots y_n) = 1$ for all $n \geq r$ if $\beta(v) = p_{K,r}(\cdot \mid v)$ for every $v \in V_{K,r}$. Otherwise there exists $c > 0$ such that $d_\beta^{(1)}(y_1 \cdots y_{r+n}) \leq e^{-cn}$ for all sufficiently large n.

We next show that every failure of continued fraction normality can be exploited by some context gambler at a positive exponential rate. The construction reuses the limiting distribution obtained in the proof that non-normality forces the entropy ratio below 1.

▶ **Theorem 20.** *If x is not continued fraction normal, then there exist parameters $K \geq 1$, $r \geq 0$, and a (K,r)-context gambler β such that $\limsup_{n\to\infty} \frac{1}{n} \log d_\beta^{(1)}(y_1 \cdots y_{r+n}) > 0$. In particular, β wins at a positive exponential rate along an infinite subsequence.*

We can now state the dichotomy in its final form.

▶ **Theorem 21** (Schnorr–Stimm dichotomy for context gamblers)**.** *Let $x \in (0,1)\setminus\mathbb{Q}$ be irrational, and for each $K \geq 1$ let $y = y^{(K)}(x)$ be its K-truncated continued fraction expansion. Then the following dichotomy holds.*

1. *If x is continued fraction normal, then for every $K \geq 1$, every $r \geq 0$, and every (K,r)-context gambler β, either*

$$d_\beta^{(1)}(y_1 \cdots y_{r+n}) = 1 \quad \textit{for all } n \geq 0,$$

or there exists $c > 0$ such that

$$d_\beta^{(1)}(y_1 \cdots y_{r+n}) \leq e^{-cn} \quad \textit{for all sufficiently large } n.$$

2. *If x is not continued fraction normal, then there exist $K \geq 1$, $r \geq 0$, and a (K,r)-context gambler β such that, for some $c > 0$,*

$$d_\beta^{(1)}(y_1 \cdots y_{r+n}) \geq e^{cn} \quad \textit{for infinitely many } n.$$

Thus the restricted context-gambler model recovers the exact continued fraction analogue of the classical Schnorr–Stimm picture: normality is precisely the point at which finite-state betting loses all positive exponential advantage.

8 Discussion and open questions

We have introduced a bounded-memory dimension theory for continued fractions with the expected core features: an entropy-rate formulation, a characterization of continued fraction normality at dimension 1, and a Schnorr–Stimm dichotomy. These results suggest several further questions.

A natural question concerns the *full-past* model. Here, the payoff rule depends on the entire past while the gambler still has only finite memory, a distinction that does not arise in the base-b setting because the one-step conditional law is constant. This leads to the following question: *Does there exist a continued fraction non-normal real number x such that, under the full-past payoff rule, no finite-state gambler wins on x at a positive exponential rate?* A positive answer would show that the model fails to capture non-normality as well, while a negative answer would say that full-past gamblers still detect every failure of normality.

A broader problem is to understand how much of the classical finite-state dimension program extends to continued fractions. In the base-b setting, one guiding theme is Lutz's *Normality–Dimension Thesis* [15], which postulates that every theorem concerning Borel normality is the dimension-1 special case of a more general theorem concerning finite-state dimension. Our results make it possible to pose the continued fraction version: *"Every theorem concerning continued fraction normality is the dimension-*1 *special case of a more general theorem concerning continued fraction finite-state dimension."* However, the continued fraction setting is less rigid because the Gauss measure is not a product measure. One basic example is selection: in the base-b setting, arithmetic-progression subsequences preserve normality, whereas in the continued fraction setting every nontrivial arithmetic-progression subsequence of a continued fraction normal number is non-normal [10]. This does not by itself refute the *Normality–Dimension Thesis* for continued fractions, but it suggest that the continued fraction setting is a natural place to test its limits.

Two concrete tests for the *Normality–Dimension Thesis* in the continued fraction setting are the following. First, Vandehey proved that nondegenerate rational matrix transformations preserve continued fraction normality [22]. Similar preservation phenomena are known for base-b normality and base-b finite-state dimension [24, 8, 20, 3]. Does the same kind of transformation also preserve continued fraction finite-state dimension? Second, sliding-block and disjoint-block definitions of continued fraction normality are known to be equivalent [17]. Since our entropy formulation is based on sliding empirical frequencies, one can ask whether the corresponding definition using disjoint blocks gives the same value. Equivalently, do the sliding-block and disjoint-block entropy rates lead to the same notion of continued fraction dimension?

A further direction is to develop a theory of continued fraction finite-state strong dimension [2]. On the gambling side, this means replacing infinitely-often unbounded capital by capital tending to infinity along the sequence. On the entropy side, the corresponding change is to replace the lower entropy rate by the upper entropy rate. In the base-b setting, known entropy-rate arguments for strong finite-state dimension use finite-state memory in an essential way, for instance to combine finitely many local betting strategies. Since context gamblers have no additional state information beyond the visible truncated context, these arguments no longer apply in the current setting. A strong-dimension theory in this setting therefore seems to require new methods.

References

1 Roy L. Adler, Michael S. Keane, and Meir Smorodinsky. A construction of a normal number for the continued fraction transformation. *Journal of Number Theory*, 13:95–105, 1981. URL: https://api.semanticscholar.org/CorpusID:122007474.

2 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.

3 Verónica Becher, Olivier Carton, and Santiago Figueira. Rauzy complexity and block entropy. *Information and Computation*, 306:105330, 2025. doi:10.1016/j.ic.2025.105330.

4 Verónica Becher and Sergio A Yuhjtman. On absolutely normal and continued fraction normal numbers. *International Mathematics Research Notices*, 2019(19):6136–6161, October 2019. doi:10.1093/imrn/rnx297.

5 C. Bourke, J. M. Hitchcock, and N. V. Vinodchandran. Entropy rates and finite-state compression. *Theoretical Computer Science*, 349(3):392–406, 2005.

6 Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley Series in Telecommunications. John Wiley & Sons, Inc., New York, 1991.

7 Jack J. Dai, James I. Lathrop, Jack H. Lutz, and Elvira Mayordomo. Finite-state dimension. *Theoretical Computer Science*, 310(1-3):1–33, 2004. doi:10.1016/S0304-3975(03)00244-5.

8 D. Doty, J. H. Lutz, and S. Nandakumar. Finite state dimension and real arithmetic. In *Proceedings of the 33rd International Colloquium on Automata, Logic and Programming*, 2006.

9 Manfred Einsiedler and Thomas Ward. *Ergodic theory*. Springer, 2013.

10 B. Heersink and J. Vandehey. Continued fraction normality is not preserved along arithmetic progressions. *Archiv der Mathematik*, 106(4):363–370, April 2016.

11 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.

12 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.

13 Marius Iosifescu and Cor Kraaikamp. *Metrical theory of continued fractions*, volume 547. Springer Science & Business Media, 2013.

14 A. Ya. Khinchin. *Continued Fractions*. Dover Publications, 1997.

15 Jack H. Lutz. Alan turing in the twenty-first century: normal numbers, randomness, and finite automata. *CCR 2012: 7th conference on computability, complexity and randomnes*, July 2012. URL: https://www.newton.ac.uk/seminar/2265/.

16 Satyadev Nandakumar. An effective ergodic theorem and some applications. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, pages 39–44, New York, NY, USA, 2008. Association for Computing Machinery. doi:10.1145/1374376.1374383.

17 Satyadev Nandakumar, Subin Pulari, Prateek Vishnoi, and Gopal Viswanathan. An analogue of pillai's theorem for continued fraction normality and an application to subsequences. *Bulletin of the London Mathematical Society*, 53(5):1414–1428, 2021. doi:10.1112/blms.12507.

18 Satyadev Nandakumar, Akhil S, and Prateek Vishnoi. Effective continued fraction dimension versus effective hausdorff dimension of reals. *ACM Trans. Comput. Theory*, 17(2), June 2025. doi:10.1145/3723323.

19 Satyadev Nandakumar and Prateek Vishnoi. On continued fraction randomness and normality. *Information and Computation*, 285:104876, 2022. doi:10.1016/j.ic.2022.104876.

20 G. Rauzy. Nombres normaux et processus déterministes. *Acta Arithmetica*, 29:211–225, 1976. doi:10.4064/aa-29-3-211-225.

21 C.-P. Schnorr and H. Stimm. Endliche Automaten und Zufallsfolgen. *Acta Informatica*, 1(4):345–359, 1972. doi:10.1007/bf00289514.

22 Joseph Vandehey. Non-trivial matrix actions preserve normality for continued fractions. *Compositio Mathematica*, 153(2):274–293, 2017. doi:10.1112/S0010437X16007740.

23 Prateek Vishnoi. Normality, Randomness and Kolmogorov Complexity of Continued Fractions. In Ding-Zhu Du, Donglei Du, Chenchen Wu, and Dachuan Xu, editors, *Theory and Applications of Models of Computation*, pages 382–392, Cham, 2022. Springer International Publishing. doi:10.1007/978-3-031-20350-3_31.

24 D. D. Wall. *Normal Sequences*. PhD thesis, University of California, Berkeley, 1949.