Santa Fe Institute Working Paper 13-09-028
arxiv.org:1309.3792 [cond-mat.stat-mech]

Exact Complexity: The Spectral Decomposition of Intrinsic Computation

James P. Crutchfield,[1,*] Christopher J. Ellison,[2,†] and Paul M. Riechers[1,‡]

[1] *Complexity Sciences Center and Department of Physics, University of California at Davis, One Shields Avenue, Davis, CA 95616*
[2] *Center for Complexity and Collective Computation, University of Wisconsin-Madison, Madison, WI 53706*

(Dated: January 16, 2016)

We give exact formulae for a wide family of complexity measures that capture the organization of hidden nonlinear processes. The spectral decomposition of operator-valued functions leads to closed-form expressions involving the full eigenvalue spectrum of the mixed-state presentation of a process's ϵ-machine causal-state dynamic. Measures include correlation functions, power spectra, past-future mutual information, transient and synchronization informations, and many others. As a result, a direct and complete analysis of intrinsic computation is now available for the temporal organization of finitary hidden Markov models and nonlinear dynamical systems with generating partitions and for the spatial organization in one-dimensional systems, including spin systems, cellular automata, and complex materials via chaotic crystallography.

Keywords: excess entropy, statistical complexity, projection operator, residual, resolvent, entropy rate, predictable information, bound information, ephemeral information

PACS numbers: 02.50.-r 89.70.+c 05.45.Tp 02.50.Ey 02.50.Ga

The emergence of organization in physical, engineered, and social systems is a fascinating and now, after half a century of active research, widely appreciated phenomenon [1–5]. Success in extending the long list of instances of emergent organization, however, is not equivalent to understanding what organization itself is. How do we say objectively that new organization has appeared? How do we measure quantitatively how organized a system has become?

Computational mechanics' answer to these questions is that a system's organization is captured in how it stores and processes information—how it computes [6]. *Intrinsic computation* was introduced two decades ago to analyze the inherent information processing in complex systems [7]: How much history does a system remember? In what architecture is that information stored? And, how does the system use it to generate future behavior?

Computational mechanics, though, is part of a long historical trajectory focused on developing a physics of information [8–10]. That nonlinear systems actively process information goes back to Kolmogorov [11], who adapted Shannon's communication theory [12] to measure the information production rate of chaotic dynamical systems. In this spirit, today computational mechanics is routinely used to determine physical and intrinsic computational properties in single-molecule dynamics [13], in complex materials [14], and even in the formation of social structure [15], to mention several recent examples.

Thus, measures of complexity are important to quantifying how organized nonlinear systems are: their randomness and their structure. Moreover, we now know that randomness and structure are intimately intertwined. One cannot be properly defined or even practically measured without the other [16, and references therein].

Measuring complexity has been a challenge: Until recently, in understanding the varieties of organization to be captured; still practically, in terms of estimating metrics from experimental data. One major reason for these challenges is that systems with emergent properties are hidden: We do not have direct access to their internal, often high-dimensional state space; we do not know a priori what the emergent patterns are. Thus, we must "reconstruct" their state space and dynamics [17–20]. Even then, when successful, reconstruction does not lead easily or directly to measures of structural complexity and intrinsic computation [7]. It gives access to what is hidden, but does not say what the mechanisms are nor how they work.

Our view of the various kinds of complexity and their measures, though, has become markedly clearer of late. There is a natural semantics of complexity in which each measure answers a specific question about a system's organization. For example:

- How random is a process? Its *entropy rate* h_μ [11].
- How much information is stored? Its *statistical com-*

* chaos@ucdavis.edu
† cellison@wisc.edu
‡ pmriechers@ucdavis.edu

plexity C_μ [7].

- How much of the future can be predicted? Its *past-future mutual information* or *excess entropy* **E** [16].
- How much information must an observer extract to know a process's hidden states? Its *transient information* **T** and *synchronization information* **S** [16].
- How much of the generated information (h_μ) affects future behavior? Its *bound information* b_μ [21].
- What's forgotten? Its *ephemeral information* ρ_μ [21].

And there are other useful measures ranging from degrees of irreversibility to quantifying model redundancy; see, for example, Ref. [22] and the proceedings in Refs. [23, 24].

Unfortunately, except in the simplest cases where expressions are known for several, to date typically measures of intrinsic computation require extensive numerical simulation and estimation. Here we answer this challenge, providing exact expressions for a process's measures in terms of its ϵ-machine. In particular, we show that the spectral decomposition of this hidden dynamic leads to closed-form expressions for complexity measures. In this way, the remaining task in analyzing intrinsic computation reduces to mathematically constructing or reliably estimating a system's ϵ-machine in the first place.

Background Our main object of study is a process $\mathcal{P}$, by which we mean the rather prosaic listing of all of a system's behaviors or realizations $\{\ldots x_{-2}, x_{-1}, x_0, x_1, \ldots\}$ and their probabilities: $\Pr(\ldots X_{-2}, X_{-1}, X_0, X_1, \ldots)$. We assume the process is stationary and ergodic and the measurement values range over a finite alphabet: $x \in \mathcal{A}$. This class describes a wide range of processes from statistical mechanical systems in equilibrium and in nonequilibrium steady states to nonlinear dynamical systems in discrete and continuous time on their attracting invariant sets.

Following Shannon and Kolmogorov, information theory gives a natural measure of a process's randomness as the uncertainty in measurement blocks: $H(L) = H[X_{0:L}]$, where H is the Shannon-Boltzmann entropy of the distribution governing the block $X_{0:L} = X_0, X_1, \ldots, X_{L-1}$. We monitor the *block entropy growth*—the average uncertainty in the next measurement X_{L-1} conditioned on knowing the preceding block $X_{0:L-1}$:

$$\begin{aligned} h_\mu(L) &= H(L) - H(L-1) \\ &= H[X_{L-1}|X_{0:L-1}] \\ &= -\left\langle \sum_{x_{L-1}\in\mathcal{A}} p(x_{L-1}) \log_2 p(x_{L-1}) \right\rangle_{\Pr(X_{0:L-1})} , \end{aligned} \tag{1}$$

where $p(x_{L-1}) = \Pr(x_{L-1}|x_{0:L-1})$. And when the limit exists, we say the process generates information at the *entropy rate*: $h_\mu = \lim_{L\to\infty} h_\mu(L)$.

Measurements, though, only indirectly reflect a system's internal organization. Computational mechanics extracts that hidden organization via the process's ϵ-machine [6], consisting of a set of recurrent *causal states* $\boldsymbol{\mathcal{S}} = \{\sigma^0, \sigma^1, \sigma^2, \ldots\}$ and transition dynamic $\{T^{(x)} : T^{(x)}_{i,j} = \Pr(x, \sigma^j|\sigma^i)\}_{x\in\mathcal{A}}$. Each causal state represents a collection of "equivalent" histories—equivalent in the sense that each history belonging to an equivalence class yields the same prediction over futures. The ϵ-machine is a system's unique, minimal-size, optimal predictor from which two key complexity measures can be directly calculated.

The entropy rate follows immediately from the ϵ-machine as the causal-state averaged transition uncertainty:

$$h_\mu = -\sum_{\sigma\in\boldsymbol{\mathcal{S}}} \Pr(\sigma) \sum_{x\in\mathcal{A}} \Pr(x|\sigma) \log_2 \Pr(x|\sigma) \ . \tag{2}$$

Here, the causal-state distribution $\Pr(\mathcal{S})$ is the stationary distribution $\langle\pi| = \langle\pi|\,T$ of the internal Markov chain governed by the row-stochastic matrix $T = \sum_{x\in\mathcal{A}} T^{(x)}$. The conditional probabilities $\Pr(x|\sigma)$ are the associated transition components in the labeled matrices $T^{(x)}_{\sigma,\sigma'}$. Note that the next state σ' is uniquely determined by knowing the current state σ and the measurement value x—a key property called *unifilarity*.

The amount of historical information the process stores also follows immediately: the *statistical complexity*, the Shannon-Boltzmann entropy of the causal-state distribution:

$$C_\mu = -\sum_{\sigma\in\boldsymbol{\mathcal{S}}} \Pr(\sigma) \log_2 \Pr(\sigma) \ , \tag{3}$$

In this way, the ϵ-machine allows one to directly determine two important properties of a system's intrinsic computation: its information generation and its storage. Since it depends only on block entropies, however, h_μ can be calculated via other presentations; though not as efficiently. For example, h_μ can be determined from Eq. (2) using any unifilar predictor, which necessarily is always larger than the ϵ-machine. Only recently was a (rather more complicated) closed-form expression discovered for the excess entropy **E** using a representation closely related to the ϵ-machine [22]. Details aside, no analogous closed-form expressions for the other complexity measures are known, including and especially those for finite-L blocks, such as $h_\mu(L)$.

Mixed-State Presentation To develop these, we shift to consider how an observer represents its knowledge of a hidden system's current state and then introduce a spectral analysis of that representation. For our uses here,

the observer has a correct model in the sense that it reproduces $\mathcal{P}$ exactly. (Any model that does we call a *presentation* of the process. There may be many.) Using this, the observer tracks a process's evolution using a distribution over the hidden states called a *mixed state* $\eta \equiv \big(\Pr(\sigma^0), \Pr(\sigma^1), \Pr(\sigma^2), \ldots\big)$. The associated random variable is denoted $\mathcal{R}$. The question is how does an observer update its knowledge (η) of the internal states as it makes measurements—$x_0, x_1, \ldots$?

If a system is in mixed state η, then the probability of seeing measurement x is: $\Pr(X = x|\mathcal{R} = \eta) = \langle\eta| T^{(x)} |\mathbf{1}\rangle$, where $\langle\eta|$ is the mixed state as a row vector and $|\mathbf{1}\rangle$ is the column vector of all 1s. This extends to measurement sequences $w = x_0 x_1 \ldots x_{L-1}$, so that if, for example, the process is in statistical equilibrium, $\Pr(w) = \langle\pi| T^{(w)} |\mathbf{1}\rangle = \langle\pi| T^{(x_0)} T^{(x_1)} \cdots T^{(x_{L-1})} |\mathbf{1}\rangle$. The mixed-state evolution induced by measurement sequence w is: $\langle\eta_{t+L}| = \langle\eta_t| T^{(w)} / \langle\eta_t| T^{(w)} |\mathbf{1}\rangle$. The set $\boldsymbol{\mathcal{R}}$ of mixed states that we use here are those induced by all allowed words $w \in \mathcal{A}^*$ from initial mixed state $\eta_0 = \pi$. For each mixed state η_{t+1} induced by symbol $x \in \mathcal{A}$, the mixed-state-to-state transition probability is: $\Pr(\eta_{t+1}, x|\eta_t) = \Pr(x|\eta_t)$. And so, by construction, using mixed states gives a unifilar presentation. We denote the associated set of transition matrices $\{W^{(x)}\}$. They and the mixed states $\boldsymbol{\mathcal{R}}$ define a process's *mixed-state presentation* (MSP), which describes how an observer's knowledge of the hidden process updates via measurements. The row-stochastic matrix $W = \sum_{x\in\mathcal{A}} W^{(x)}$ governs the evolution of the probability distribution over allowed mixed states.

The use of mixed states is originally due to Blackwell [25], who expressed the entropy rate h_μ as an integral of a (then uncomputable) measure over the mixed-state state space $\boldsymbol{\mathcal{R}}$. Although we focus here on the finite mixed-state case for simplicity, it is instructive to see in the general case the complicatedness revealed in a process using the mixed-state presentation: e.g., Figs. 17(a)-(c) of Ref. [26]. The Supplementary Materials give the detailed calculations for examples in the finite case.

Mixed states allow one to derive an efficient expression for the finite-L entropy-rate estimates of Eq. (1):

$$h_\mu(L) = H\big[X_{L-1}|\big(\mathcal{R}_{L-1}|\mathcal{R}_0 = \pi\big)\big] \ . \tag{4}$$

This says that one need only update the initial distribution over mixed states (with all probability density on $\eta_0 = \pi$) to the distribution at time L by tracking powers W^L of the internal transition dynamic of the MSP and not tracking, for that matter, an exponentially growing number of intervening sequences $\{x^L\}$. (This depends critically on the MSP's unifilarity.) That is, using the MSP reduces the original exponential computational complexity of estimating the entropy rate to polynomial time in L. Finally, and more to the present task, the mixed-state simplification is the main lead to an exact, closed-form analysis of complexity measures, achieved by combining the MSP with a spectral decomposition of the mixed-state evolution as governed by W^L.

Spectral Decomposition State distribution evolution involves iterating the transition dynamic W^L—that is, taking powers of a row-stochastic square matrix. As is well known, functions of a diagonalizable matrix can often be carried out efficiently by operating on its eigenvalues. More generally, using the Cauchy integral formula for operator-valued functions [27] and given W's eigenvalues $\Lambda_W \equiv \{\lambda \in \mathbb{C} : \det(\lambda I - W) = 0\}$, we find that W^L's spectral decomposition is:

$$W^L = \sum_{\substack{\lambda\in\Lambda_W \\ \lambda\neq 0}} \lambda^L W_\lambda \left\{ I + \sum_{N=1}^{\nu_\lambda - 1} \binom{L}{N} \left(\lambda^{-1} W - I\right)^N \right\} + [0 \in \Lambda_W] \left\{ \delta_{L,0} W_0 + \sum_{N=1}^{\nu_0 - 1} \delta_{L,N} W_0 W^N \right\}, \tag{5}$$

where $[0 \in \Lambda_W]$ is the Iverson bracket (unity when $\lambda = 0$ is an eigenvalue, vanishing otherwise), $\delta_{i,j}$ is the Kronecker delta function, and ν_λ is the size of the largest Jordan block associated with λ: $\nu_\lambda \leq 1 + a_\lambda - g_\lambda$, where g_λ and a_λ are λ's geometric (subspace dimension) and algebraic (order in the characteristic polynomial) multiplicities, respectively. The matrices $\{W_\lambda\}$ are a mutually orthogonal set of projection operators given by the residues of W's resolvent:

$$W_\lambda = \tfrac{1}{2\pi i} \oint_{C_\lambda} (zI - W)^{-1} dz \ , \tag{6}$$

a counterclockwise integral around singular point λ.

For simplicity here, consider only Ws that are diagonalizable. In this case: $g_\lambda = a_\lambda$ and Eq. (5) simplifies to $W^L = \sum_{\lambda\in\Lambda_W} \lambda^L W_\lambda$, where the projection operators reduce to $W_\lambda = \prod_{\substack{\zeta\in\Lambda_W \\ \zeta\neq\lambda}} (W - \zeta I) / (\lambda - \zeta)$. Thus, the only L-dependent operation in forming W^L is simply exponentiating its eigenvalues. The powers determine all of a process's properties, both transient (finite-L) and asymptotic. Moreover, when $a_\lambda = 1$, there are further simplifications. In particular, $W_\lambda = |\boldsymbol{\lambda}\rangle \langle\boldsymbol{\lambda}| / \langle\boldsymbol{\lambda}|\boldsymbol{\lambda}\rangle$, where $|\boldsymbol{\lambda}\rangle$ and $\langle\boldsymbol{\lambda}|$ are the right and left eigenvectors, respectively, of W associated with λ. Contrary to the Hermitian case of quantum mechanics, the left eigenvectors are not simply the conjugate transpose of the right eigenvectors.

Closed-Form Complexities Forming the mixed-state presentation of process's ϵ-machine, its spectral decomposition leads directly to analytic, closed-form expressions for many complexity measures—here we present

formulae only for $h_\mu(L)$, $\mathbf{E}$, $\mathbf{S}$, and $\mathbf{T}$. Similar expressions for correlation functions and power spectra, partition functions, b_μ, r_μ, and others are presented elsewhere.

Starting from its mixed-state expression in Eq. (4) for the length-L entropy-rate approximates $h_\mu(L)$, we find the closed-form expression:

$$\begin{aligned} h_\mu(L) &= \langle \delta_\pi | \, W^{L-1} \, | H(W^{\mathcal{A}}) \rangle \\ &= \sum_{\lambda \in \Lambda_W} \lambda^{L-1} \, \langle \delta_\pi | \, W_\lambda \, | H(W^{\mathcal{A}}) \rangle \ , \end{aligned} \tag{7}$$

where δ_π is the distribution with all probability density on the MSP's unique start state—the mixed state corresponding to the ϵ-machine's equilibrium distribution π [28]. In addition, $|H(W^{\mathcal{A}})\rangle$ is a column vector of transition uncertainties from each allowed mixed state η:

$$|H(W^{\mathcal{A}})\rangle = -\sum_{\eta \in \boldsymbol{\mathcal{R}}} |\delta_\eta\rangle \sum_{x \in \mathcal{A}} \langle \delta_\eta | \, W^{(x)} \, |\mathbf{1}\rangle \log_2 \langle \delta_\eta | \, W^{(x)} \, |\mathbf{1}\rangle \ .$$

Taking $L \to \infty$, one finds the entropy rate (cf. Eq. (2)):

$$h_\mu = \langle \delta_\pi | \, W_1 \, | H(W^{\mathcal{A}}) \rangle = \langle \pi_W | H(W^{\mathcal{A}}) \rangle \ ,$$

where π_W is the stationary distribution over the MSP.

Let's turn to analyze the past-future mutual information, the excess entropy $\mathbf{E} = I[X_{-\infty:0}; X_{0:\infty}]$: the information from the past that reduces uncertainty in the future. In general, $\mathbf{E}$ is not the statistical complexity C_μ, which is the information from the past that must be stored in order make optimal predictions about the future. Although Eq. (3) makes it clear that the stored information C_μ is immediately calculable from the ϵ-machine, $\mathbf{E}$ is substantially less direct. To see this, recall that the excess entropy has an equivalent definition—$\mathbf{E} = \sum_{L=1}^{\infty} [h_\mu(L) - h_\mu]$—to which we can apply Eq. (7), obtaining:

$$\mathbf{E} = \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{1-\lambda} \, \langle \delta_\pi | \, W_\lambda | H(W^{\mathcal{A}}) \rangle \ . \tag{8}$$

This should be compared to the only previously known general closed-form, which uses a process and its time-reversal [22, 29]: $\mathbf{E} = I[\mathcal{S}^-; \mathcal{S}^+]$, where $\mathcal{S}^-$ and $\mathcal{S}^+$ are the causal states of the reverse-time and forward-time ϵ-machines, respectively. Intriguingly, it appears that Eq. (8)—the spectral expression from the forward process—captures aspects of the reverse-time process. In short, the MSP's transient structure encapsulates the relevant information in the recurrent causal states of the reverse process. This suggests that there is a transformation between transient states of a forward process and recurrent states of its reverse process. (We explore this in a sequel.)

How does an observer come to know a hidden process's internal state? We monitor this via the *average state uncertainty* having seen all length-L words [16]:

$$\begin{aligned} \mathcal{H}(L) &= -\sum_{w \in \mathcal{A}^L} \Pr(w) \sum_{\sigma \in \boldsymbol{\mathcal{S}}} \Pr(\sigma|w) \log_2 \Pr(\sigma|w) \\ &= \sum_{\eta \in \boldsymbol{\mathcal{R}}} \Pr(\mathcal{R}_L = \eta | \mathcal{R}_0 = \pi) H[\eta] \ , \end{aligned}$$

where the last line is the mixed-state version with $H[\eta]$ being the presentation-state uncertainty specified by the mixed state η. (Although recurrent causal states are induced by infinite histories, a partial history w, as above, starting from the stationary distribution π, can induce a nonpeaked distribution over causal states: $0 \leq \Pr(\sigma|w) \leq 1$.) Applying the spectral decomposition yields, for diagonalizable W:

$$\mathcal{H}(L) = \langle \delta_\pi | \, W^L \, | H[\eta] \rangle = \sum_{\lambda \in \Lambda_W} \lambda^L \, \langle \delta_\pi | \, W_\lambda \, | H[\eta] \rangle \ ,$$

where $|H[\eta]\rangle$ is the column vector of state-distribution uncertainties for each allowed mixed state $\eta \in \boldsymbol{\mathcal{R}}$.

The observer is *synchronized* when the state uncertainty vanishes: $\mathcal{H}(L) = 0$. The total amount of state information, then, that an observer must extract to become synchronized is the *synchronization information* [16] $\mathbf{S} \equiv \sum_{L=0}^{\infty} \mathcal{H}(L)$. Applying the above spectral decomposition results in the following closed-form expression:

$$\mathbf{S} = \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{1-\lambda} \, \langle \delta_\pi | \, W_\lambda | H[\eta] \rangle \ . \tag{9}$$

This form makes it clear that mixed states and the transitions between them capture fundamental properties of the underlying process. For example, rewriting C_μ as the entropy of the initial mixed state—$\langle \delta_\pi | H[\eta] \rangle$—reinforces this observation. A related measure, the presentation-dependent synchronization information [30], diverges when the presentation is not synchronizable. Note also the close similarity between the excess entropy formula Eq. (8) above and Eq. (9). The only difference is that they average different informational quantities—the transition or the state uncertainties, respectively.

Although there are a number of additional complexity measures, as we discussed above, the final example we present is the *transient information* $\mathbf{T}$ [16]. It measures the amount of information one must extract from observations so that the block entropy converges to its linear asymptote: $\mathbf{T} \equiv \sum_{L=1}^{\infty} L \, [h_\mu(L) - h_\mu]$. The spectral decomposition readily yields the following closed-form ex-

pression:

$$\mathbf{T} = \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda|<1}} \frac{1}{(1-\lambda)^2} \langle \delta_\pi | \, W_\lambda | H(W^{\mathcal{A}}) \rangle \ .$$

This form reveals the close relationship between transient information and excess entropy: they differ only in the eigenvalue weighting.

Discussion There are a number of comments to make at this point to draw out the results' usefulness and importance. A process's structural complexity is not controlled by only the first spectral gap—the difference between the maximum and next eigenvalue. Rather, the *entire* spectrum is implicated in calculating the complexity measures. In terms of the temporal dynamics, all subspaces of the underlying causal-state process contribute. Naturally, there will be cases in which only some subspaces dominate, but as the expressions show this is not the general case. In addition, there is much structural information to be extracted from the projection operators W_λ, such as the dimension (geometric multiplicity) of the associated subspaces on which they act. This, in turn, gives the number of active degrees of freedom for the constituent subprocesses. As a result, we see that complexity measures capture inherently different properties, far beyond pairwise correlations and power spectra. Specifically, they measure transient, finite-time, and time asymptotic information processing, properties that rely on all-way correlations. The result gives one of the most thorough quantitative views of a process' intrinsic computation [7].

Although their derivations have not been laid out, the formulae as given are immediately usable. The Supplementary Materials provide calculations of complexity measures for several examples, including those for processes generated by typical unifilar hidden Markov models, nonlinear dynamical systems, cellular automata, and materials that form close-packed structures, as determined experimentally from X-ray diffraction spectra. And, a sequel demonstrates how to analyze the intrinsic computation in the stochastic switching dynamics of neural ion channels [31, 32].

One of the more direct physical consequences of computational mechanics is that a system's organization is synonymous with how it stores and transforms information—a view complementary to that physics takes in terms of energy storage and transduction. In short, how a system is organized is how it computes. The theory just introduced grounds this information processing view practically and mathematically, giving quantitative and exact analysis of how hidden processes are organized. And, it should be contrasted with Kolmogorov-Chaitin complexity analysis [33]. For both stochastic and deterministic chaotic processes, the Kolmogorov-Chaitin complexity is (i) dominated by randomness and (ii) uncomputable. The theory here could not be more different: A wide variety of distinct kinds of information storage and processing are identified and they are exactly calculable.

The authors thank Ryan James and Dowman Varn for helpful discussions. JPC is a member of the Santa Fe Institute External Faculty. This material is based upon work supported by, or in part by, the U. S. Army Research Laboratory and the U. S. Army Research Office under contracts W911NF-13-1-0390, W911NF-13-1-0340, and W911NF-12-1-0234 and by a grant from the John Templeton Foundation.

[1] C. Domb, M. S. Green, and J. L. Lebowitz, editors. *Phase Transitions and Critical Phenomena*, volume 1-20. Academic Press, 1972–2001.

[2] M. C. Cross and P. C. Hohenberg. *Rev. Mod. Phys.*, 65(3):851–1112, 1993.

[3] S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraulas, and E. Bonabeau. *Self-Organization in Biological Systems*. Princeton University Press, New York, 2003.

[4] M. Newman, A.-L. Barabasi, and D. J. Watts. *The Structure and Dynamics of Networks*. Princeton University Press, Princeton, New Jersey, 2006.

[5] D. Helbing, editor. *Social Self-Organization*. Springer, New York, 2012.

[6] J. P. Crutchfield. *Nature Physics*, 8(January):17–24, 2012.

[7] J. P. Crutchfield and K. Young. *Phys. Rev. Lett.*, 63:105–108, 1989.

[8] W. Weaver. *Am. Sci.*, 36:536, 1948.

[9] L. Brillouin. *Science and Information Theory*. Academic Press, New York, 1956.

[10] R. Landauer. *IBM J. Res. Develop.*, 5(3):183–191, 1961.

[11] A. N. Kolmogorov. *Dokl. Akad. Nauk. SSSR*, 124:754, 1959. (Russian) Math. Rev. vol. 21, no. 2035b.

[12] C. E. Shannon. *Bell Sys. Tech. J.*, 27:379–423, 623–656, 1948.

[13] C.-B. Li, H. Yang, and T. Komatsuzaki. *Proc. Natl. Acad. Sci. USA*, 105:536–541, 2008.

[14] D. P. Varn, G. S. Canright, and J. P. Crutchfield. *Acta. Cryst. Sec. A*, 69(2):197–206, 2013.

[15] D. Darmon, J. Sylvester, M. Girvan, and W. Rand. page arxiv.org: 1306.6111, 2013.

[16] J. P. Crutchfield and D. P. Feldman. *CHAOS*, 13(1):25–54, 2003.

[17] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw. *Phys. Rev. Let.*, 45:712, 1980.

[18] F. Takens. In D. A. Rand and L. S. Young, editors, *Symposium on Dynamical Systems and Turbulence*, volume 898, page 366, Berlin, 1981. Springer-Verlag.

[19] J. P. Crutchfield and B. S. McNamara. *Complex Systems*, 1:417 – 452, 1987.

[20] J. D. Farmer and J. Sidorowitch. *Phys. Rev. Lett.*, 59:366, 1987.

[21] R. G. James, C. J. Ellison, and J. P. Crutchfield. *CHAOS*, 21(3):037109, 2011.

[22] J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney. *Phys. Rev. Lett.*, 103(9):094101, 2009.

[23] B. Marcus, K. Petersen, and T. Weissman, editors. *Entropy of Hidden Markov Process and Connections to Dynamical Systems*, volume 385 of *Lecture Notes Series*. London Mathematical Society, 2011.

[24] J. P. Crutchfield and J. Machta. *CHAOS*, 21(3):037101, 2011.

[25] D. Blackwell. volume 28, pages 13–20, Publishing House of the Czechoslovak Academy of Sciences, Prague, 1957.

[26] J. P. Crutchfield. *Physica D*, 75:11–54, 1994.

[27] S. Hassani. *Mathematical Physics*. Springer, New York, 1999.

[28] To be clear, let's review: π is the stationary distribution over the ϵ-machine's states. It also happens to be—as a distribution over $\mathbf{S}$ —a mixed state. π does not denote the initial distribution over the mixed states, δ_π does, as it is the distribution peaked at the mixed-state π. The use of δ_π is consistent with our use of δ_η for other distributions over the MSP that are peaked at the mixed-state η.

[29] C. J. Ellison, J. R. Mahoney, and J. P. Crutchfield. *J. Stat. Phys.*, 136(6):1005–1034, 2009.

[30] J. P. Crutchfield, C. J. Ellison, J. R. Mahoney, and R. G. James. *CHAOS*, 20(3):037105, 2010.

[31] P. Kienker. *Proc. Roy. Soc. Lond. B*, 236:269–309, 1989.

[32] W. J. Bruno, J. Yang, and J. E. Pearson. *PNAS*, 102(18):6326–6331, 2005.

[33] M. Li and P. M. B. Vitanyi. *An Introduction to Kolmogorov Complexity and its Applications*. Springer-Verlag, New York, 1993.

SUPPLEMENTARY MATERIALS

The following sections provide detailed calculations of the mixed-state presentations and the measures of intrinsic computation discussed in the main text using the closed-form expressions there. We analyze first a prototypic strictly sofic system, the Even Process, and then give results for the symbolic dynamics of the Tent Map at a Misiurewicz parameter, the spacetime domain for elementary cellular automaton rule 22, and finally the chaotic crystallographic structure of a close-packed polytypic material—Zinc Sulfide—as determined from experimental X-ray diffractograms.

Appendix A: Even Process, a prototype sofic system

Consider the Even Process, the stochastic process generated over the two-symbol alphabet $\mathcal{A} = \{\square, \triangle\}$ by the hidden Markov model (HMM) $\mathcal{M}$ of Fig. 1. Though finitely specified, being a strictly sofic system it produces sequences that have arbitrarily long correlations; e.g., sequences containing blocks $\square(\triangle\triangle)^k\square$, for $k = 0, 1, 2, \ldots$.

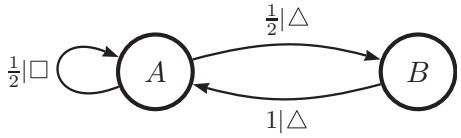

FIG. 1. An HMM $\mathcal{M}$ (in this case, the ϵ-machine) that generates the Even Process.

Ordering the causal states alphabetically, $\mathcal{M}$ has symbol-labeled transition matrices:

$$T^{(\square)} = \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } T^{(\triangle)} = \begin{bmatrix} 0 & 1/2 \\ 1 & 0 \end{bmatrix} .$$

From the state-transition matrix:

$$T = \begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix} , \tag{A1}$$

we find the stationary distribution:

$$\pi = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix} . \tag{A2}$$

The mixed-state presentation $\mathcal{M}_{\text{msp}}$ gives the dynamics over $\mathcal{M}$'s state distributions, starting from the stationary distribution $\langle\pi|$, induced by observed words $w \in \mathcal{A}^*$. $\mathcal{M}_{\text{msp}}$ can be algorithmically generated from $\mathcal{M}$. A state transition diagram for $\mathcal{M}_{\text{msp}}$ is given in Fig. 2.

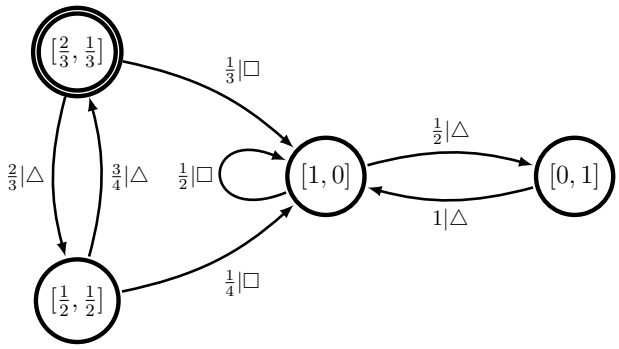

FIG. 2. $\mathcal{M}$'s mixed-state presentation $\mathcal{M}_{\text{msp}}$. The doubly-circled state denotes the start state $\eta_0 = \pi$. Inside each mixed state is a vector $[\Pr(A), \Pr(B)]$ which describes the induced distribution over the states of $\mathcal{M}$.

$\mathcal{M}_{\text{msp}}$ has state-transition matrix:

$$W = \sum_{x \in \mathcal{A}} W^{(x)} = \begin{bmatrix} 0 & 2/3 & 1/3 & 0 \\ 3/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1 & 0 \end{bmatrix} . \tag{A3}$$

Solving $\det(\lambda I - W) = 0$ gives W's eigenvalues:

$$\Lambda_W = \left\{1, \tfrac{\sqrt{2}}{2}, -\tfrac{\sqrt{2}}{2}, -\tfrac{1}{2}\right\} .$$

For each, we find the corresponding projection operator W_λ via:

$$W_\lambda = \prod_{\substack{\zeta \in \Lambda_W \\ \zeta \neq \lambda}} \frac{W - \zeta I}{\lambda - \zeta} ,$$

obtaining:

$$W_1 = \begin{bmatrix} 0 & 0 & 2/3 & 1/3 \\ 0 & 0 & 2/3 & 1/3 \\ 0 & 0 & 2/3 & 1/3 \\ 0 & 0 & 2/3 & 1/3 \end{bmatrix} ,$$

$$W_{\sqrt{2}/2} = \begin{bmatrix} 1/2 & \sqrt{2}/3 & -(2+\sqrt{2})/6 & -(1+\sqrt{2})/6 \\ 3\sqrt{2}/8 & 1/2 & -(1+\sqrt{2})/4 & -(2+\sqrt{2})/8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} ,$$

$$W_{-\sqrt{2}/2} = \begin{bmatrix} 1/2 & -\sqrt{2}/3 & -(2-\sqrt{2})/6 & -(1-\sqrt{2})/6 \\ -3\sqrt{2}/8 & 1/2 & -(1-\sqrt{2})/4 & -(2-\sqrt{2})/8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} ,$$

and

$$W_{-1/2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1/6 & 1/6 \\ 0 & 0 & 1/3 & -1/3 \\ 0 & 0 & -2/3 & 2/3 \end{bmatrix} .$$

Note that $W_1 = |\mathbf{1}\rangle\langle\pi_W|$, which is always the case for an ergodic process.

We construct $\langle\delta_\pi|$ by placing all of the initial mass at $\mathcal{M}_{\text{msp}}$'s start state, representing the stationary distribution π over the original presentation $\mathcal{M}$:

$$\langle\delta_\pi| = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} .$$

Table I collects together the spectral quantities of W necessary for the exact calculation of the various complexity measures.

Different measures of complexity track the evolution of different types of information in (or about) the system. The entropy of transitioning from the various states of uncertainty is given by the ket $|H(W^{\mathcal{A}})\rangle$, whereas the internal entropy of the states of uncertainty themselves is given by the ket $|H[\eta]\rangle$. From the labeled transition matrices of the mixed-state presentation, we find:

$$|H(W^{\mathcal{A}})\rangle = \begin{bmatrix} \log_2(3) - 2/3 \\ 2 - \frac{3}{4}\log_2(3) \\ 1 \\ 0 \end{bmatrix} .$$

And from the mixed states themselves,

$$\sum_{\eta\in\boldsymbol{\mathcal{R}}} \eta\,|\delta_\eta\rangle = \begin{bmatrix} (2/3\,,\ 1/3) \\ (1/2\,,\ 1/2) \\ (1\,,\ 0) \\ (0\,,\ 1) \end{bmatrix} ,$$

we have

$$|H[\eta]\rangle = \begin{bmatrix} \log_2(3) - 2/3 \\ 1 \\ 0 \\ 0 \end{bmatrix} .$$

Hence, the finite-L entropy-rate convergence for $L \geq 1$ is given by:

$$\begin{aligned} h_\mu(L) &= \sum_{\lambda\in\Lambda_W} \lambda^{L-1}\,\langle\delta_\pi|\,W_\lambda|H(W^{\mathcal{A}})\rangle \\ &= \begin{cases} \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1}\left(-\frac{\sqrt{2}}{2}\log_2(3) + \sqrt{2}\right) & \text{for even } L \\ \frac{2}{3} + \left(\frac{\sqrt{2}}{2}\right)^{L-1}\left(\log_2(3) - \frac{4}{3}\right) & \text{for odd } L \end{cases} . \end{aligned}$$

This function is shown in Fig. 3. Similarly, the average

λ	ν_λ	$\langle\delta_\pi\|\,W_\lambda$			
1	1	$[0$	0	$\frac{2}{3}$	$\frac{1}{3}]$
$\sqrt{2}/2$	1	$[\frac{1}{2}$	$\frac{\sqrt{2}}{3}$	$\frac{-2-\sqrt{2}}{6}$	$\frac{-1-\sqrt{2}}{6}]$
$-\sqrt{2}/2$	1	$[\frac{1}{2}$	$\frac{-\sqrt{2}}{3}$	$\frac{-2+\sqrt{2}}{6}$	$\frac{-1+\sqrt{2}}{6}]$
$-1/2$	1	$[0$	0	0	$0]$

TABLE I. Useful spectral quantities for the Even Process.

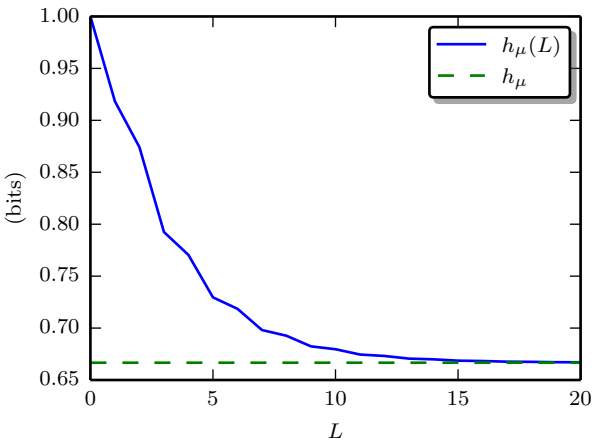

FIG. 3. The entropy-rate convergence $h_\mu(L)$ for the Even Process as a function of L, imposing the boundary condition that $h_\mu(0) = \log_2|\mathcal{A}| = 1$.

state uncertainty after L observations is given by:

$$\begin{aligned} \mathcal{H}(L) &= \sum_{\lambda\in\Lambda_W} \lambda^L\,\langle\delta_\pi|\,W_\lambda|H[\eta]\rangle \\ &= \left(\tfrac{\sqrt{2}}{2}\right)^L \left[\tfrac{1}{2}\log_2(3) - \tfrac{1}{3} + \tfrac{\sqrt{2}}{3} \right. \\ &\qquad \left. +(-1)^L\left(\tfrac{1}{2}\log_2(3) - \tfrac{1}{3} - \tfrac{\sqrt{2}}{3}\right)\right] \\ &= \begin{cases} (\log_2(3) - 2/3)\left(\frac{\sqrt{2}}{2}\right)^L & \text{for even } L \\ \frac{2\sqrt{2}}{3}\left(\frac{\sqrt{2}}{2}\right)^L & \text{for odd } L \end{cases} . \end{aligned}$$

For the scalar complexity measures, we find:

$$\begin{aligned} h_\mu &= 2/3 \text{ bits per step,} \\ C_\mu &= \log_2(3) - \tfrac{2}{3} \text{ bits,} \\ \mathbf{E} &= \log_2(3) - \tfrac{2}{3} \text{ bits,} \\ \mathbf{T} &= 2\log_2(3) \text{ bits-symbols, and} \\ \mathbf{S} &= 2\log_2(3) \text{ bits.} \end{aligned}$$

Appendix B: Misiurewicz Tent Map: Intrinsic Computation in a Continuous-state Chaotic Dynamical System

The Tent Map of the unit interval:

$$x_{n+1} = a \cdot \min\{x_n, 1 - x_n\} \ ,$$

for $x_0 \in [0, 1]$, is a well studied chaotic dynamical system. We set the parameter a to one of the so-called Misiurewicz values:

$$\begin{aligned} a &= \alpha + \tfrac{2}{3\alpha} \\ &\approx 1.76929235 \ , \end{aligned}$$

where $\alpha = \sqrt[3]{\sqrt{19/27} + 1}$. At this setting, the iterate of the map's maximum ($x_c = 1/2$) becomes period-1 after the third iterate. For a detailed analysis of this map see Ref. [A1]. Here, we use the ϵ-machine found there for the binary symbolic dynamics observed with a generating partition that divides the interval at x_c; see Fig. 4.

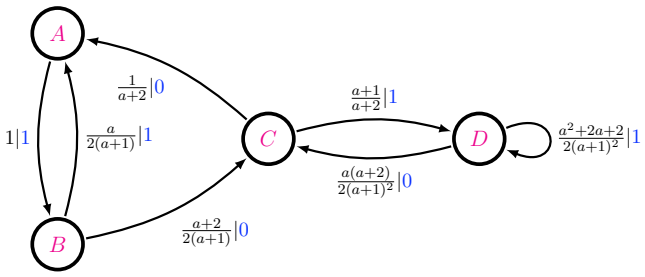

FIG. 4. The ϵ-machine $\mathcal{M}$ that captures the binary symbolic dynamics at the chosen Misiurewicz parameter. The transitions are labeled $p|x$ where p is the transition probability and $x \in \{0, 1\}$ is the emitted symbol. Adapted from Ref. [A1] with permission.

$\mathcal{M}$ has the two-symbol alphabet $\mathcal{A} = \{0, 1\}$ with the corresponding symbol-labeled transition matrices:

$$T^{(0)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{a+2}{2(a+1)} & 0 \\ \frac{1}{a+2} & 0 & 0 & 0 \\ 0 & 0 & \frac{a(a+2)}{2(a+1)^2} & 0 \end{bmatrix} \text{ and}$$

$$T^{(1)} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{a}{2(a+1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{a+1}{a+2} \\ 0 & 0 & 0 & \frac{a^2+2a+2}{2(a+1)^2} \end{bmatrix} .$$

We find the stationary distribution from the state-transition matrix T:

$$\langle \pi | = \begin{bmatrix} \pi_A & \pi_B & \pi_C & \pi_D \end{bmatrix} \ , \tag{B1}$$

with

$$\pi_A = \pi_B = \frac{2(a + a^2)}{2 + 14a + 14a^2 + 3a^3} \tag{B2}$$

$$\pi_C = \frac{4a + 4a^2 + a^3}{2 + 14a + 14a^2 + 3a^3} \tag{B3}$$

$$\pi_D = \frac{2(1 + 3a + 3a^2 + a^3)}{2 + 14a + 14a^2 + 3a^3} \ . \tag{B4}$$

So,

$$\begin{aligned} \langle \eta_\varnothing | &= \langle \pi | \\ &\propto \begin{bmatrix} 2a(a+1) & 2a(a+1) & a(a+2)^2 & 2(a+1)^3 \end{bmatrix} \ , \\ \langle \eta_0 | &\propto \langle \eta_\varnothing | \, T^{(0)} \\ &\propto \begin{bmatrix} 1 & 0 & a+2 & 0 \end{bmatrix} \ , \\ \langle \eta_1 | &\propto \langle \eta_\varnothing | \, T^{(1)} \\ &\propto \begin{bmatrix} a^2 & 2a(a+1) & 0 & 2(a+1)^3 \end{bmatrix} \ , \\ \langle \eta_{00} | &\propto \langle \eta_0 | \, T^{(0)} \\ &\propto \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \ , \\ \langle \eta_{01} | &\propto \langle \eta_0 | \, T^{(1)} \\ &\propto \begin{bmatrix} 0 & 1 & 0 & a+1 \end{bmatrix} \ , \\ \langle \eta_{10} | &\propto \langle \eta_1 | \, T^{(0)} \\ &\propto \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \ , \\ \langle \eta_{11} | &\propto \langle \eta_1 | \, T^{(1)} \\ &\propto \begin{bmatrix} a^2 & a^2 & 0 & (a+1)(a^2+2a+2) \end{bmatrix} \ , \\ \langle \eta_{001} | &\propto \langle \eta_{00} | \, T^{(1)} \\ &\propto \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \ , \\ \langle \eta_{011} | &\propto \langle \eta_{01} | \, T^{(1)} \\ &\propto \begin{bmatrix} a & 0 & 0 & a^2+2a+2 \end{bmatrix} \ , \end{aligned}$$

and

$$\begin{aligned} \langle \eta_{101} | &\propto \langle \eta_{10} | \, T^{(1)} \\ &\propto \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \ , \end{aligned}$$

give all of the distinct mixed states, labeled in subscripts by the first shortest word that induces the mixed state from the initial mixed state $\eta_\varnothing = \pi$.

The mixed-state presentation is shown in Fig. 5. Keying off the graph's topology, we order the mixed states as:

$$\boldsymbol{\mathcal{R}} = \{\pi, \ \eta_0, \ \eta_{01}, \eta_{011}, \ \eta_1, \eta_{11}, \ \overset{A}{\eta_{00}}, \overset{B}{\eta_{001}}, \overset{C}{\eta_{10}}, \overset{D}{\eta_{101}}\},$$

which is the ordering we use for the transition matrix and the bras and kets. We put the ϵ-machine's recurrent state names—A, B, C, and D—above the last four mixed states since they are isomorphic. That is, the recurrent

states of the mixed-state presentation (of the ϵ-machine) are effectively the recurrent states of the ϵ-machine endowed with peaked distributions that uniquely identify themselves among the causal states.

The transition probabilities between mixed states can be calculated via:

$$W_{\eta_\varnothing \to \eta_0} = \frac{\langle \eta_\varnothing | \, T^{(0)} \, | \mathbf{1} \rangle}{\langle \eta_\varnothing | \, \mathbf{1} \rangle} = \frac{a(a+2)(a+3)}{3a^3+14a^2+14a+2} \approx 0.364704$$

$$W_{\eta_\varnothing \to \eta_1} = \frac{\langle \eta_\varnothing | \, T^{(1)} \, | \mathbf{1} \rangle}{\langle \eta_\varnothing | \, \mathbf{1} \rangle} = \frac{2a^3+9a^2+8a+2}{3a^3+14a^2+14a+2} \approx 0.635296$$

$$W_{\eta_0 \to \eta_{00}} = \frac{\langle \eta_0 | \, T^{(0)} \, | \mathbf{1} \rangle}{\langle \eta_0 | \, \mathbf{1} \rangle} = \frac{1}{a+3} \approx 0.209675$$

and so on, to determine the transition matrix for the mixed-state presentation:

$$W = \sum_{x \in \mathcal{A}} W^{(x)} = \begin{bmatrix}
0 & \frac{a(a+2)(a+3)}{3a^3+14a^2+14a+2} & 0 & 0 & \frac{2a^3+9a^2+8a+2}{3a^3+14a^2+14a+2} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{a+2}{a+3} & 0 & 0 & 0 & \frac{1}{a+3} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\
0 & 0 & \frac{a}{a+1} & 0 & 0 & 0 & 0 & 0 & \frac{1}{a+1} & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{a^3+5a^2+4a+2}{2a^3+9a^2+8a+2} & 0 & 0 & \frac{a(a+2)^2}{2a^3+9a^2+8a+2} & 0 \\
0 & 0 & 0 & 0 & \frac{a^3+6a^2+4a+4}{2(a^3+5a^2+4a+2)} & 0 & 0 & 0 & \frac{a(a+2)^2}{2(a^3+5a^2+4a+2)} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{a}{2(a+1)} & 0 & \frac{a+2}{2(a+1)} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{a+2} & 0 & 0 & \frac{a+1}{a+2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{a(a+2)}{2(a+1)^2} & \frac{a^2+2a+2}{2(a+1)^2}
\end{bmatrix}.$$

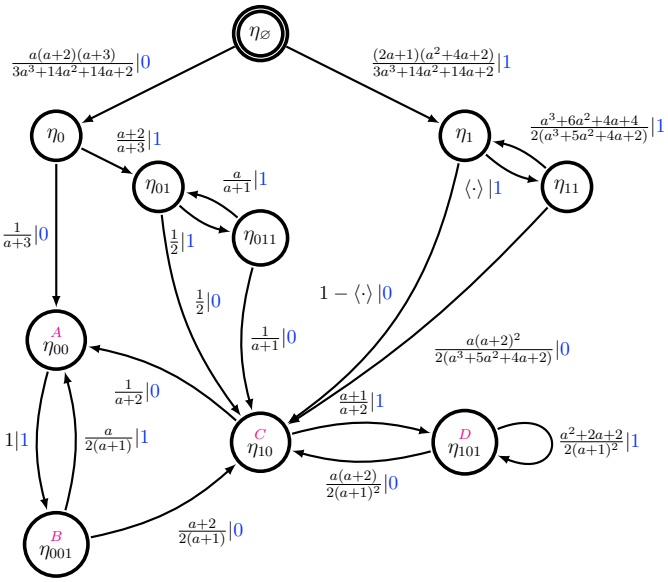

FIG. 5. The mixed-state presentation $\mathcal{M}_{\text{msp}}$ of the symbolic dynamics of the Tent Map at a Misiurewicz parameter. Due to space constraints we represent $W_{\eta_1 \to \eta_{11}} = \frac{a^3+5a^2+4a+2}{(2a+1)(a^2+4a+2)}$ as $\langle \cdot \rangle$ and $W_{\eta_1 \to \eta_{10}} = \frac{a(a+2)^2}{(2a+1)(a^2+4a+2)}$ as $1 - \langle \cdot \rangle$.

The mixed states themselves can be cast as the column vector:

$$\sum_{\eta \in \boldsymbol{\mathcal{R}}} \eta \, |\delta_\eta\rangle = \begin{bmatrix}
\frac{2(a+1)}{3a^3+14a^2+14a+2}\left(a, a, \frac{a(a+2)^2}{2(a+1)}, (a+1)^2\right) \\
\frac{1}{a+3}\left(1, 0, a+2, 0\right) \\
\frac{1}{a+2}\left(0, 1, 0, a+1\right) \\
\frac{1}{(a+1)(a+2)}\left(a, 0, 0, \frac{2(a+1)^2}{a}\right) \\
\frac{1}{2a^3+9a^2+8a+2}\left(a^2, 2a(a+1), 0, 2(a+1)^3\right) \\
\frac{1}{a^3+5a^2+4a+2}\left(a^2, a^2, 0, \frac{2(a+1)^3}{a}\right) \\
\left(1, 0, 0, 0\right) \\
\left(0, 1, 0, 0\right) \\
\left(0, 0, 1, 0\right) \\
\left(0, 0, 0, 1\right)
\end{bmatrix},$$

from which we obtain the internal entropy of the mixed states. This is perhaps best left represented as:

$$|H[\eta]\rangle = \sum_{\eta \in \boldsymbol{\mathcal{R}}} |\delta_\eta\rangle \times - \sum_{\sigma \in \boldsymbol{\mathcal{S}}} \langle \delta_\sigma | \, \eta \rangle \log_2 \langle \delta_\sigma | \, \eta \rangle$$

$$= \begin{bmatrix} H[\pi] \\ H[\eta_0] \\ H[\eta_{01}] \\ H[\eta_{011}] \\ H[\eta_1] \\ H[\eta_{11}] \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \approx \begin{bmatrix} 1.731517488310359 \\ 0.740860503264943 \\ 0.834641915284059 \\ 0.656560846029285 \\ 0.970187276943296 \\ 0.942259650948105 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ bits.}$$

Evidently, the internal entropy of the recurrent mixed states (the last four rows) is zero—as expected of the recurrent mixed states of the ϵ-machine. From $C_\mu = \langle \delta_\pi | H[\eta] \rangle$, the statistical complexity of the process is $C_\mu = H[\pi]$.

Let Q be the 6×6 substochastic matrix of transitions among the ordered set of transient mixed states $\boldsymbol{\mathcal{R}}_Q = \{\pi,\ \eta_0,\ \eta_{01}, \eta_{011},\ \eta_1, \eta_{11}\}$. Then, for some 6×4 matrix B, we write:

$$W = \begin{bmatrix} Q & B \\ 0 & T \end{bmatrix},$$

where the '0' just above is the 4×6 matrix of all zeros. We see that to obtain the synchronization information, we need only powers of Q. B and T are irrelevant to calculating $\mathbf{S}$, once Q has been obtained:

$$\begin{aligned} \mathbf{S} &= \sum_{L=0}^{\infty} \langle \delta_\pi | W^L | H[\eta] \rangle \\ &= \sum_{L=0}^{\infty} \begin{bmatrix} \langle \delta_\pi | & 0 \end{bmatrix} \begin{bmatrix} Q^L & (\cdot) \\ 0 & T^L \end{bmatrix} \begin{bmatrix} | H[\eta \in \boldsymbol{\mathcal{R}}_Q] \rangle \\ 0 \end{bmatrix} \\ &= \sum_{L=0}^{\infty} \langle \delta_\pi | Q^L | H[\eta \in \boldsymbol{\mathcal{R}}_Q] \rangle , \end{aligned}$$

where the zeros that appear above inherit the appropriate dimensions for matrix multiplication. And, we reuse the notation $\langle \delta_\pi |$ to refer to the shortened 1×6 row-vector with the last four zeros removed. We are now ready to obtain the synchronization information via the spectral decomposition of Q^L.

From:

$$\begin{aligned} \det(\lambda I - Q) &= \lambda^2 \left(\lambda^2 - \tfrac{a}{2(a+1)}\right) \left(\lambda^2 - \tfrac{a^3+6a^2+4a+4}{2(2a^3+9a^2+8a+2)}\right) \\ &= \lambda^2 \left(\lambda^2 - \tfrac{a}{2(a+1)}\right)^2 , \end{aligned}$$

we obtain Q's eigenvalues:

$$\Lambda_Q = \left\{0, \pm\sqrt{\tfrac{a}{2(a+1)}}\right\} .$$

Since the index of the zero eigenvalue is greater than one ($\nu_0 = 2$, since $a_0 = 2$ and $g_0 = 1$), Q is not diagonalizable. Nevertheless, since all eigenvalues besides the zero eigenvalue have index equal to unity, we can calculate powers of Q by the slightly more general spectral decomposition given in Eq. (5):

$$\begin{aligned} Q^L &= \left\{ \sum_{\lambda \in \Lambda_Q} \lambda^L Q_\lambda \right\} + \sum_{N=1}^{\nu_0 - 1} \delta_{L,N} Q_0 Q^N \\ &= \left\{ \sum_{\lambda \in \Lambda_Q} \lambda^L Q_\lambda \right\} + \delta_{L,1} Q_0 Q . \end{aligned} \tag{B5}$$

Using Eq. (B5), the synchronization information becomes:

$$\begin{aligned} \mathbf{S} = &\langle \delta_\pi | Q_0 Q | H[\eta \in \boldsymbol{\mathcal{R}}_Q] \rangle \\ &+ \sum_{\lambda \in \Lambda_Q} \frac{1}{1-\lambda} \langle \delta_\pi | Q_\lambda | H[\eta \in \boldsymbol{\mathcal{R}}_Q] \rangle . \end{aligned} \tag{B6}$$

Since the two eigenvalues besides the zero eigenvalue have index equal to unity, their projection operators can be obtained via:

$$\begin{aligned} Q_\lambda &= \left(\frac{Q}{\lambda}\right)^{\nu_0} \prod_{\substack{\zeta \in \Lambda_Q \setminus \{0\} \\ \zeta \neq \lambda}} \frac{Q - \zeta I}{\lambda - \zeta} \\ &= \left(\frac{Q}{\lambda}\right)^2 \frac{Q + \lambda I}{2\lambda} \\ &= \tfrac{a+1}{a} Q^2 \left(\lambda^{-1} Q + I\right) \end{aligned}$$

for each $\lambda \in \left\{\pm\sqrt{\tfrac{a}{2(a+1)}}\right\}$. Then, since the projection operators must sum to the identity matrix, Q_0 can be obtained via:

$$\begin{aligned} Q_0 &= I - \sum_{\lambda \in \Lambda_Q \setminus \{0\}} Q_\lambda \\ &= I - \tfrac{a+1}{a} Q^2 \left(\sqrt{\tfrac{2(a+1)}{a}} Q - \sqrt{\tfrac{2(a+1)}{a}} Q + 2I \right) \\ &= I - \tfrac{2(a+1)}{a} Q^2 . \end{aligned}$$

Preparing to calculate $\mathbf{S}$, we find:

$$\sum_{\lambda \in \Lambda_Q \setminus \{0\}} \frac{1}{1-\lambda} \langle \delta_\pi | Q_\lambda$$

$$= \tfrac{a+1}{a} \langle \delta_\pi | \left\{ \left(\frac{1}{\sqrt{\frac{a}{2(a+1)}} - \frac{a}{2(a+1)}} \right) Q^3 + \left(\frac{1}{1 - \sqrt{\frac{a}{2(a+1)}}} \right) Q^2 + \left(\frac{-1}{\sqrt{\frac{a}{2(a+1)}} + \frac{a}{2(a+1)}} \right) Q^3 + \left(\frac{1}{1 + \sqrt{\frac{a}{2(a+1)}}} \right) Q^2 \right\}$$
$$= \tfrac{a+1}{a} \langle \delta_\pi | \left\{ \left(\frac{2}{1 - \frac{a}{2(a+1)}} \right) Q^3 + \left(\frac{2}{1 - \frac{a}{2(a+1)}} \right) Q^2 \right\}$$
$$= \tfrac{4(a+1)^2}{a(a+2)} \langle \delta_\pi | \left(Q^3 + Q^2 \right) .$$

Hence:

$$\sum_{\lambda \in \Lambda_Q} \frac{1}{1-\lambda} \langle \delta_\pi | Q_\lambda$$
$$= \langle \delta_\pi | \left\{ I + \tfrac{2(a+1)}{a} \left(\tfrac{2(a+1)}{a+2} - 1 \right) Q^2 + \tfrac{4(a+1)^2}{a(a+2)} Q^3 \right\}$$
$$= \langle \delta_\pi | \left(I + \tfrac{2(a+1)}{a+2} Q^2 + \tfrac{4(a+1)^2}{a(a+2)} Q^3 \right)$$

and:

$$\langle \delta_\pi | Q_0 Q + \sum_{\lambda \in \Lambda_Q} \frac{1}{1-\lambda} \langle \delta_\pi | Q_\lambda$$
$$= \langle \delta_\pi | \left(Q - \tfrac{2(a+1)}{a} Q^3 \right) + \langle \delta_\pi | \left(I + \tfrac{2(a+1)}{a+2} Q^2 + \tfrac{4(a+1)^2}{a(a+2)} Q^3 \right)$$
$$= \langle \delta_\pi | \left\{ I + Q + \tfrac{2(a+1)}{a+2} \left(Q^2 + Q^3 \right) \right\} . \qquad \text{(B7)}$$

From (B6) and (B7), we obtain the synchronization information,

$$\mathbf{S} = \langle \delta_\pi | \left(I + Q + \tfrac{2(a+1)}{a+2} \left(Q^2 + Q^3 \right) \right) | H[\eta \in \boldsymbol{\mathcal{R}}_Q] \rangle ,$$

which can be written more explicitly as:

$$\mathbf{S} = H[\pi] + \tfrac{1}{3a^3+14a^2+14a+2} \Big\{ a(a+2)(a+3) H[\eta_0]$$
$$+ a(a+1)(a+2) \Big(2H[\eta_{01}] + H[\eta_{011}] \Big)$$
$$+ \tfrac{2(a+1)}{a+2} \Big(\tfrac{a+1}{a} (a^3 + 6a^2 + 4a + 4) H[\eta_1]$$
$$+ (a^3 + 5a^2 + 4a + 2) H[\eta_{11}] \Big) \Big\}$$
$$\approx 3.880442712215985 \text{ bits.}$$

Appendix C: Intrinsic Computation in Spacetime

Here, we calculate the informational properties of one of elementary cellular automaton (ECA) rule 22's domains—its dominant spacetime invariant sets. See Ref. [A2] for a discussion of spacetime domain and particle analysis and related structures for ECA 54. We mention only the minimal prerequisites necessary for analyzing spacetime intrinsic computation.

ECA 22 generates a series of 1D spatial configurations. The domain is a collection of spacetime patches consisting of binary values $\{0, 1\}$ at each site that is dynamically invariant under spatial and temporal shifts. It is described by the ϵ-machine shown in Fig. 6 over the four-letter alphabet $\mathcal{A} = \{0_s, 1_s, 0_t, 1_t\}$, where the site values are subscripted by s when the value is seen when making a *spacelike* move and by t when the value is seen when making a *timelike* move. Thus, the ϵ-machine describes the set of binary strings observed when taking all (right) spacelike and (positive) timelike moves across the spacetime lattice.

$\mathcal{M}$ has the symbol-labeled transition matrices:

$$T^{(0_s)} = \begin{bmatrix} 0 & v & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & v & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & v & 0 & 0 & 0 & 0 & 0 & 0 \\ v/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & v & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & v \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & v & 0 & 0 \end{bmatrix} ,$$

$$T^{(1_s)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ v/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & v/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & v & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & v & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} ,$$

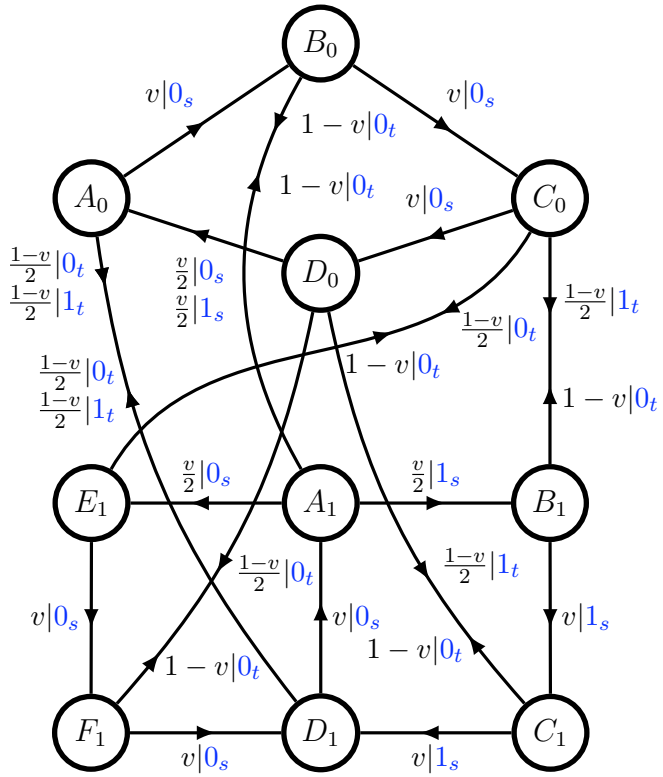

FIG. 6. The ϵ-machine $\mathcal{M}$ describing the spacetime configuration patches generated by ECA 22 that are spacetime shift invariant. When leaving a state, spacelike (s) and timelike (t) moves are made with probability dependent on the average "velocity" $v \in (0,1)$. Here, $v = 0$ corresponds to a strictly timelike sequence of moves and $v = 1$ corresponds to a strictly spacelike sequence of moves. Once a spacelike or timelike move has been chosen, if the topology of the ϵ-machine does not completely determine the next site value, then the site values (0 or 1) occur with equal probability. The product of these form the transition probabilities listed on the edges. Note that the state-to-state transitions D_0 to A_0, D_1 to A_0, and A_0 to D_1 consist of two parallel transitions on 0 and 1, though they are depicted with a single edge in the diagram. Those edges do have a pair of labels as appropriate.

$$T^{(0_t)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1-v}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-v & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1-v}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1-v}{2} \\ 0 & 1-v & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-v & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-v & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1-v}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-v & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-v & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ and}$$

$$T^{(1_t)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1-v}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-v}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1-v}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1-v}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The state-transition matrix $T = \sum_{x \in \mathcal{A}} T^{(x)}$ is:

$$T = \begin{bmatrix} 0 & v & 0 & 0 & 0 & 0 & 0 & 1-v & 0 & 0 \\ 0 & 0 & v & 0 & 1-v & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & v & 0 & \frac{1-v}{2} & 0 & 0 & \frac{1-v}{2} & 0 \\ v & 0 & 0 & 0 & 0 & 0 & \frac{1-v}{2} & 0 & 0 & \frac{1-v}{2} \\ 0 & 1-v & 0 & 0 & 0 & v/2 & 0 & 0 & v/2 & 0 \\ 0 & 0 & 1-v & 0 & 0 & 0 & v & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-v & 0 & 0 & 0 & v & 0 & 0 \\ 1-v & 0 & 0 & 0 & v & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-v & 0 & 0 & 0 & 0 & 0 & 0 & v \\ 0 & 0 & 0 & 1-v & 0 & 0 & 0 & v & 0 & 0 \end{bmatrix}.$$

And, from it we find the stationary distribution:

$$\langle \pi | = \tfrac{1}{16} \begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 1 & 1 & 2 & 1 & 1 \end{bmatrix} . \qquad \text{(C1)}$$

However, at the extremes of $v = 0$ and $v = 1$ the ϵ-machine breaks apart into an ensemble of subprocesses. We analyze several subprocesses here, comparing the complexity measures of path ensembles in space versus those in time.

1. Timelike complexity

For $v = 0$, one of the strictly timelike subprocesses collapses down to the noisy period-2 process shown in Fig. 7.

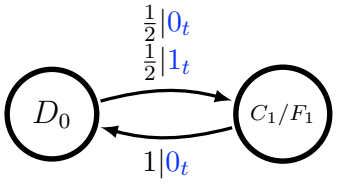

FIG. 7. The ϵ-machine of one of the strictly timelike subprocesses that appears at $v = 0$.

For this strictly timelike subprocess, we obtain the MSP shown in Fig. 8, where the mixed states are labeled with their corresponding distribution $[\Pr(D_0), \Pr(C_1/F_1)]$.

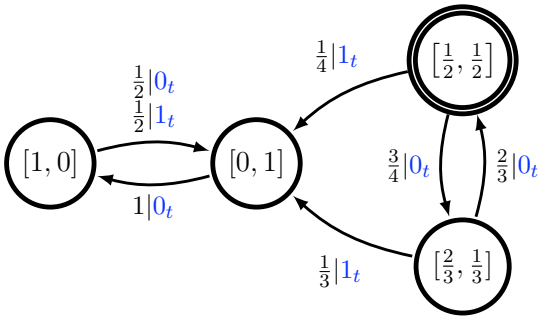

FIG. 8. The MSP of the strictly timelike subprocess shown in Fig. 7.

If we order the set of mixed states as:

$$\boldsymbol{\mathcal{R}} = \{\pi = (\tfrac{1}{2}, \tfrac{1}{2}),\ \eta_0 = (\tfrac{2}{3}, \tfrac{1}{3}),\ \eta_1 = (0,1), \eta_{10} = (1,0)\},$$

then the mixed-state presentation has state-transition matrix:

$$W = \sum_{x \in \mathcal{A}} W^{(x)} = \begin{bmatrix} 0 & 3/4 & 1/4 & 0 \\ 2/3 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} . \tag{C2}$$

Solving $\det(\lambda I - W) = 0$ gives W's eigenvalues:

$$\Lambda_W = \left\{1,\ -1,\ \sqrt{\tfrac{1}{2}},\ -\sqrt{\tfrac{1}{2}}\right\} .$$

For each, we find the corresponding projection operator W_λ via:

$$W_\lambda = \prod_{\substack{\zeta \in \Lambda_W \\ \zeta \neq \lambda}} \frac{W - \zeta I}{\lambda - \zeta} ,$$

obtaining:

$$\begin{aligned} W_1 &= -\tfrac{1}{2} I - \tfrac{1}{2} W + W^2 + W^3 \\ &= \begin{bmatrix} 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{bmatrix} , \end{aligned}$$

$$\begin{aligned} W_{-1} &= -\tfrac{1}{2} I + \tfrac{1}{2} W + W^2 - W^3 \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1/6 & 1/6 \\ 0 & 0 & 1/2 & -1/2 \\ 0 & 0 & -1/2 & 1/2 \end{bmatrix} , \end{aligned}$$

$$W_{\sqrt{2}/2} = I - W^2 + \sqrt{2}\left(W - W^3\right)$$

$$= \begin{bmatrix} 1/2 & 3\sqrt{2}/8 & -(2+\sqrt{2})/8 & -(1+\sqrt{2})/4 \\ \sqrt{2}/3 & 1/2 & -(1+\sqrt{2})/6 & -(2+\sqrt{2})/6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} ,$$

and:

$$W_{-\sqrt{2}/2} = I - W^2 - \sqrt{2}\left(W - W^3\right)$$

$$= \begin{bmatrix} 1/2 & -3\sqrt{2}/8 & -(2-\sqrt{2})/8 & -(1-\sqrt{2})/4 \\ -\sqrt{2}/3 & 1/2 & -(1-\sqrt{2})/6 & -(2-\sqrt{2})/6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} .$$

Note that $W_1 = |\mathbf{1}\rangle\langle\pi_W|$, again, since the timelike subprocess is ergodic.

We construct $\langle\delta_\pi|$ by placing all of the initial mass at $\mathcal{M}_{\text{msp}}$'s start state, representing the stationary distribution π over the original presentation $\mathcal{M}$:

$$\langle\delta_\pi| = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} .$$

Different measures of complexity track the evolution of different types of information in (or about) the system. The entropy of transitioning from the various states of uncertainty is given by the ket $|H(W^{\mathcal{A}})\rangle$, whereas the internal entropy of the states of uncertainty themselves is given by the ket $|H[\eta]\rangle$. From the labeled transition matrices of the mixed-state presentation, we find:

$$|H(W^{\mathcal{A}})\rangle = \begin{bmatrix} 2 - \frac{3}{4}\log_2(3) \\ \log_2(3) - 2/3 \\ 0 \\ 1 \end{bmatrix} .$$

And from the mixed states themselves,

$$\sum_{\eta \in \boldsymbol{\mathcal{R}}} \eta\, |\delta_\eta\rangle = \begin{bmatrix} (1/2\,,\ 1/2) \\ (2/3\,,\ 1/3) \\ (0\,,\ 1) \\ (1\,,\ 0) \end{bmatrix} ,$$

we have:

$$|H[\eta]\rangle = \begin{bmatrix} 1 \\ \log_2(3) - 2/3 \\ 0 \\ 0 \end{bmatrix} .$$

As a step in calculating $\mathbf{E}$, $\mathbf{S}$, and $\mathbf{T}$ we find:

$$\sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{1-\lambda} \langle\delta_\pi| W_\lambda = \begin{bmatrix} 2 & \frac{3}{2} & -\frac{3}{2} & -2 \end{bmatrix} , \text{ and}$$

$$\sum_{\substack{\lambda\in\Lambda_W\\ |\lambda|<1}} \frac{1}{(1-\lambda)^2}\,\langle\delta_\pi|\,W_\lambda = \begin{bmatrix}6 & 6 & -5 & -7\end{bmatrix}.$$

Hence, for the scalar complexity measures of the strictly timelike subprocess, we find:

$$\begin{aligned} h_\mu &= 1/2 \text{ bit per step,}\\ C_\mu &= 1 \text{ bit,}\\ \mathbf{E} &= 1 \text{ bit,}\\ \mathbf{T} &= 1+\tfrac{3}{2}\log_2(3) \text{ bits-symbols, and}\\ \mathbf{S} &= 1+\tfrac{3}{2}\log_2(3) \text{ bits.}\end{aligned}$$

2. Spacelike complexity

We just considered the complexity measures of the ECA 22 domain for one of the strictly timelike subprocesses at $v=0$. At the other velocity extreme of $v=1$, one of the strictly spacelike domain subprocesses is similar: the noisy period-4 process shown in Fig. 9.

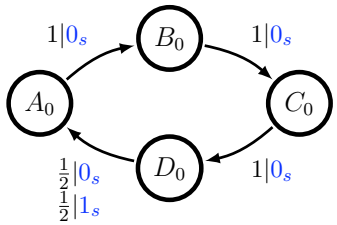

FIG. 9. ϵ-Machine of one of the strictly spacelike subprocesses at $v=1$.

For this strictly spacelike subprocess, we obtain the MSP shown in Fig. 10, where each mixed state is represented by a label corresponding to a particular state distribution $[\Pr(A_0),\Pr(B_0),\Pr(C_0),\Pr(D_0)]$.

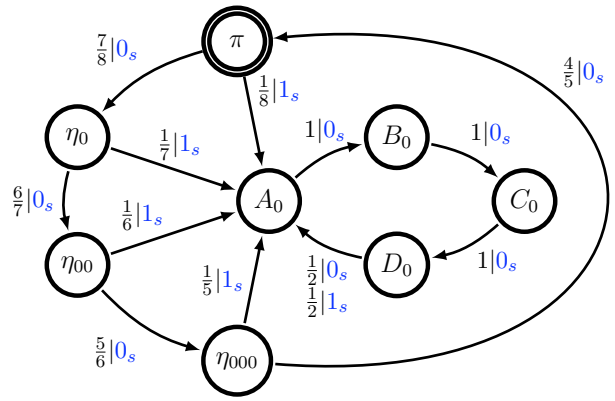

FIG. 10. MSP of the strictly spacelike subprocess of Fig. 9.

The mixed states induced from the stationary distribution by observing sequences can be cast into a column vector:

$$\sum_{\eta\in\boldsymbol{\mathcal{R}}}\eta\,|\delta_\eta\rangle = \begin{bmatrix}\pi\\ \eta_0\\ \eta_{00}\\ \eta_{000}\\ A_0\\ B_0\\ C_0\\ D_0\end{bmatrix} = \begin{bmatrix}(1/4, & 1/4, & 1/4, & 1/4)\\ (1/7, & 2/7, & 2/7, & 2/7)\\ (1/6, & 1/6, & 1/3, & 1/3)\\ (1/5, & 1/5, & 1/5, & 2/5)\\ (1, & 0, & 0, & 0)\\ (0, & 1, & 0, & 0)\\ (0, & 0, & 1, & 0)\\ (0, & 0, & 0, & 1)\end{bmatrix}. \quad \text{(C3)}$$

Using the same ordering of mixed states as in Eq. (C3), the mixed-state presentation has state-transition matrix:

$$W = \sum_{x\in\mathcal{A}} W^{(x)} = \begin{bmatrix} 0 & 7/8 & 0 & 0 & 1/8 & 0 & 0 & 0\\ 0 & 0 & 6/7 & 0 & 1/7 & 0 & 0 & 0\\ 0 & 0 & 0 & 5/6 & 1/6 & 0 & 0 & 0\\ 4/5 & 0 & 0 & 0 & 1/5 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\end{bmatrix}.$$

From Eq. (C3), we have the internal entropy of each mixed state:

$$|H[\eta]\rangle = \begin{bmatrix}2\\ \log_2(7)-6/3\\ \log_2(3)+1/3\\ \log_2(5)-2/5\\ 0\\ 0\\ 0\\ 0\end{bmatrix}.$$

From Fig. 10, we obtain the entropy of transitioning from each mixed state:

$$|H(W^{\mathcal{A}})\rangle = \begin{bmatrix}3-\frac{7}{8}\log_2(7)\\ \log_2(7)-\frac{6}{7}\log_2(3)-6/7\\ 1+\log_2(3)-\frac{5}{6}\log_2(5)\\ \log_2(5)-8/5\\ 0\\ 0\\ 0\\ 1\end{bmatrix}.$$

Solving $\det(\lambda I - W)=0$ gives W's eigenvalues, the fourth roots of unity and the fourth roots of $\frac{1}{2}$:

$$\Lambda_W = \left\{1,\,-1,\,i,\,-i,\,\sqrt[4]{\tfrac{1}{2}},\,-\sqrt[4]{\tfrac{1}{2}},\,i\sqrt[4]{\tfrac{1}{2}},\,-i\sqrt[4]{\tfrac{1}{2}}\right\}.$$

We obtain the projection operators $\{W_\lambda\}$ as in Eq. (6) via the residues of the W's resolvent around each eigen-

value in Λ_W, which are also the isolated poles of W's resolvent. In fact, we only obtain the first row $\{\langle\delta_\pi| W_\lambda\}$ of each projection operator, since finding the entire resolvent and set of projection matrices is superfluous to our immediate goal.

Employing a complex variable z, matrix inversion gives the first row of the resolvent matrix:

$$\langle\delta_\pi|(zI-W)^{-1} = \frac{1}{8(z^4-1/2)}\left[8z^3 \;\; 7z^2 \;\; 6z \;\; 5 \;\; \tfrac{z^3}{z-1} \;\; \tfrac{z^2}{z-1} \;\; \tfrac{z}{z-1} \;\; \tfrac{1}{z-1}\right] . \tag{C4}$$

The poles of each element of $\langle\delta_\pi|\,(zI-W)^{-1}$ are immediately evident: The first four elements only have poles at the four fourth roots of $1/2$. The last four entries have five poles—the four fourth roots of $1/2$ and at z equal to unity.

Since all of the poles are simple (W being diagonalizable), the projection operators can be most easily obtained by the residue algorithm for simple poles:

$$\begin{aligned} W_\lambda &= \mathrm{Res}\left((zI-W)^{-1},\, z\to\lambda\right) \\ &= \lim_{z\to\lambda}(z-\lambda)(zI-W)^{-1}, \end{aligned}$$

where the residues are taken element-wise. As a simple consequence:

$$\begin{aligned} \langle\delta_\pi|\,W_\lambda &= \mathrm{Res}\left(\langle\delta_\pi|\,(zI-W)^{-1},\, z\to\lambda\right) \\ &= \lim_{z\to\lambda}(z-\lambda)\,\langle\delta_\pi|\,(zI-W)^{-1}. \end{aligned}$$

Hence, we immediately see that the first row of all projection operators associated with the three roots of unity, besides unity itself, are row vectors of all zeros:

$$\begin{aligned} \langle\delta_\pi|\,W_{-1} &= \begin{bmatrix}0&0&0&0&0&0&0&0\end{bmatrix} \\ \langle\delta_\pi|\,W_{i} &= \begin{bmatrix}0&0&0&0&0&0&0&0\end{bmatrix} \\ \langle\delta_\pi|\,W_{-i} &= \begin{bmatrix}0&0&0&0&0&0&0&0\end{bmatrix}. \end{aligned}$$

Moreover, the first row of the projection operator associated with unity, which is also identifiable with the stationary distribution over the mixed-state presentation, is easily found to be:

$$\langle\delta_\pi|\,W_1 = \begin{bmatrix}0&0&0&0&\frac{1}{4}&\frac{1}{4}&\frac{1}{4}&\frac{1}{4}\end{bmatrix}.$$

The remaining four projection operators are all associated with eigenvalues λ such that $\lambda^4 = 1/2$. To obtain the remaining residues of Eq. (C4), we note that:

$$\frac{z-\lambda}{z^4-\lambda^4} = \frac{1}{z^3+\lambda z^2+\lambda^2 z+\lambda^3} ,$$

so that:

$$\lim_{z\to\lambda}\left(\frac{z-\lambda}{z^4-\lambda^4}\right) = \frac{1}{4}\lambda^{-3}.$$

The first row of the remaining four projection operators is thus:

$$\langle\delta_\pi|W_\lambda = \frac{1}{8}\left[2 \;\; \tfrac{7}{4\lambda} \;\; \tfrac{3}{2\lambda^2} \;\; \tfrac{5}{4\lambda^3} \;\; \tfrac{-1}{1-\lambda} \;\; \tfrac{-1}{\lambda(1-\lambda)} \;\; \tfrac{-1}{\lambda^2(1-\lambda)} \;\; \tfrac{-1}{\lambda^2(1-\lambda)}\right] ,$$

for $\lambda^4 = 1/2$.

It is then straightforward to calculate the complexity measures:

$h_\mu = 1/4$ bit per step,
$C_\mu = 2$ bits,
$\mathbf{E} = 2$ bit,
$\mathbf{T} = \frac{5}{2} + \frac{7}{4}\log_2(7) + \frac{5}{4}\log_2(5) + \frac{3}{2}\log_2(3)$ bits-symbols,
and
$\mathbf{S} = \frac{5}{2} + \frac{7}{4}\log_2(7) + \frac{5}{4}\log_2(5) + \frac{3}{2}\log_2(3)$ bits.

Since $\mathbf{E} = C_\mu$, each the timelike and spacelike subprocess shares all information that is stored from their *past* with their *future* via their *present*. In other words, the subprocesses have no crypticity: $\chi = C_\mu - \mathbf{E} = 0$.

We conclude that the strictly spacelike subprocess is less random (via h_μ), stores more information (via C_μ), shares more information with the future (via $\mathbf{E}$), and is more difficult to synchronize to (via $\mathbf{S}$ and $\mathbf{T}$) than the strictly timelike subprocess. One sees, rather explicitly now, that different facets of complexity express themselves more or less prominently along different spacetime paths within the same spacetime domain.

Appendix D: Zinc Sulfide: Intrinsic Spatial Computation in Polytypic Materials

As a final example we analyze the intrinsic computation in the spatial organization of the polytypic, closed-packed material Zinc Sulfide. Using experimentally measure X-ray diffraction spectra, Ref. [A3] extracted the ϵ-machine shown in Fig. 11. This is for the sample called SK135 there.

$\mathcal{M}$ has the two-symbol alphabet $\mathcal{A} = \{0, 1\}$ with the

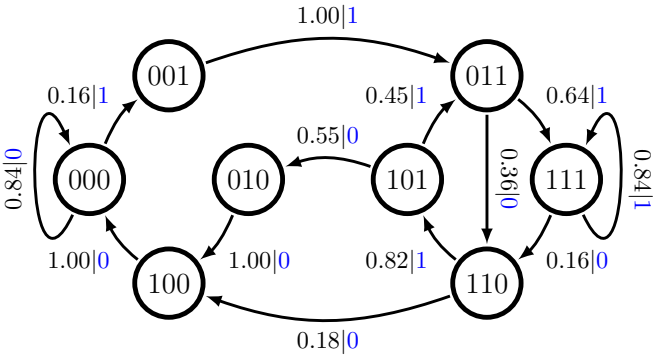

FIG. 11. ϵ-Machine $\mathcal{M}$ that captures the spatial organization in sample SK135 of Zinc Sulfide, expressed in the Hägg notation. Repeated here from Ref. [A3] with permission.

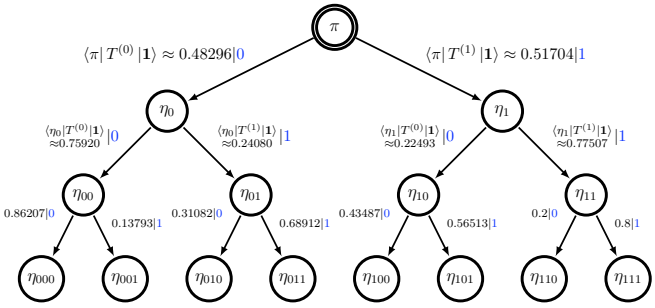

FIG. 12. The tree-like transition structure among the transient mixed states. The transition structure among the recurrent states (those states in the bottom row) is isomorphic to the recurrent structure of Fig. 11.

corresponding symbol-labeled transition matrices:

$$T^{(0)} = \begin{bmatrix} 0.84 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.36 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.55 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.18 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.16 & 0 \end{bmatrix} \text{ and}$$

$$T^{(1)} = \begin{bmatrix} 0 & 0.16 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.64 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.45 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.82 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.84 \end{bmatrix} .$$

We find the stationary distribution from the state-transition matrix T:

$$\langle \pi | \approx \begin{bmatrix} 0.32 & 0.05 & 0.04 & 0.08 & 0.05 & 0.07 & 0.08 & 0.32 \end{bmatrix} .$$

The mixed-state presentation $\mathcal{M}_{\text{msp}}$ gives the dynamics induced by observed symbols over $\mathcal{M}$'s state distributions, starting from the stationary distribution π.

Since $\mathcal{M}$ has finite Markov order of $R = 3$, the mixed-state presentation is a depth-3 tree of nonrecurrent transient states that feed into the recurrent states of $\mathcal{M}$, as shown in Fig. 12. Algebraically, this translates to the addition of a zero eigenvalue with index $\nu_0 = 3$. Hence, W is nondiagonalizable, although $a_\lambda = g_\lambda$ for all $\lambda \neq 0$. Eq. (5) then implies that:

$$W^L = \left\{ \sum_{\lambda \in \Lambda_W} \lambda^L W_\lambda \right\} + \sum_{N=1}^{\nu_0 - 1} \delta_{L,N} W_0 W^N . \qquad \text{(D1)}$$

λ	ν_λ	$\langle \delta_\pi \vert W_\lambda$
1	1	$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \dots 0 & 0 & 0.32 & 0.05 & 0.04 \\ \dots 0.08 & 0.05 & 0.07 & 0.08 & 0.32 \end{bmatrix}$
0	3	$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \dots 0 & 0 & -0.32 & -0.05 & -0.04 \\ \dots -0.08 & -0.05 & -0.07 & -0.08 & -0.32 \end{bmatrix}$
$\zeta \in \Lambda_W^{\setminus\{0,1\}}$	1	$\langle \mathbf{0} \vert$

TABLE II. Useful spectral quantities for the ZnS polytype analysis. For compactness we define $\Lambda_W^{\setminus\{0,1\}} \equiv \Lambda_W \setminus \{0, 1\}$ to be the set of W's eigenvalues that are not zero or unity. None of the projection operators associated with these other eigenvalues overlap with $\langle \delta_\pi |$. Moreover, note that $\langle \delta_\pi | W_0 = \langle \delta_\pi | (I - W_1)$ and $\langle \delta_\pi | W_1 = \langle \delta_\pi | \mathbf{1} \rangle \langle \pi_W | = \langle \pi_W |$.

Moreover, the complexity measures involve $\langle \delta_\pi | W^L$ and thus $\langle \delta_\pi | W_\lambda$. Importantly, in this example, $\langle \delta_\pi | W_\lambda = \langle \mathbf{0} |$ for all $0 < |\lambda| < 1$. In particular, $\langle \delta_\pi | W_0 = \langle \delta_\pi | (I - W_1)$ and $\langle \delta_\pi | W_1 = \langle \delta_\pi | \mathbf{1} \rangle \langle \pi_W | = \langle \pi_W |$. This is shown explicitly in Table II .

Different complexity measures track the evolution of different types of information in (or about) the system. The entropy of transitioning from the various states of uncertainty is given by the ket $|H(W^{\mathcal{A}})\rangle$, whereas the internal entropy of the states of uncertainty themselves is given by the ket $|H[\eta]\rangle$. From the labeled transition

matrices of the mixed state presentation, we find:

$$|H(W^{\mathcal{A}})\rangle \approx \begin{bmatrix} H_2(0.48296) \\ H_2(0.24080) \\ H_2(0.22493) \\ H_2(0.13793) \\ H_2(0.31082) \\ H_2(0.43487) \\ H_2(0.20000) \\ H_2(0.16) \\ 0 \\ 0 \\ H_2(0.36) \\ 0 \\ H_2(0.45) \\ H_2(0.18) \\ H_2(0.16) \end{bmatrix} \approx \begin{bmatrix} 0.9992 \\ 0.7964 \\ 0.7691 \\ 0.5788 \\ 0.8941 \\ 0.9877 \\ 0.7219 \\ 0.6343 \\ 0 \\ 0 \\ 0.9427 \\ 0 \\ 0.9928 \\ 0.6801 \\ 0.6343 \end{bmatrix},$$

where $H_2(q)$ is the binary entropy function, $H_2(q) \equiv -q\log_2(q) - (1-q)\log_2(1-q)$, and quantities associated with the transient states are colored blue. And, from the mixed states themselves:

$$\sum_{\eta \in \boldsymbol{\mathcal{R}}} \eta \, |\delta_\eta\rangle$$

$$\approx \begin{bmatrix} (0.32 & 0.05 & 0.04 & 0.08 & 0.05 & 0.07 & 0.08 & 0.32\ ) \\ (\ 0.65 & 0 & 0.07 & 0 & 0.1 & 0 & 0.17 & 0\ ) \\ (\ 0 & 0.1 & 0 & 0.16 & 0 & 0.13 & 0 & 0.62\ ) \\ (\ 0.86 & 0 & 0 & 0 & 0.14 & 0 & 0 & 0\ ) \\ (\ 0 & 0.43 & 0 & 0 & 0 & 0.57 & 0 & 0\ ) \\ (\ 0 & 0 & 0.31 & 0 & 0 & 0 & 0.69 & 0\ ) \\ (\ 0 & 0 & 0 & 0.2 & 0 & 0 & 0 & 0.8\ ) \\ (1 & 0 & 0 & 0 & 0 & 0 & 0 & 0) \\ (0 & 1 & 0 & 0 & 0 & 0 & 0 & 0) \\ (0 & 0 & 1 & 0 & 0 & 0 & 0 & 0) \\ (0 & 0 & 0 & 1 & 0 & 0 & 0 & 0) \\ (0 & 0 & 0 & 0 & 1 & 0 & 0 & 0) \\ (0 & 0 & 0 & 0 & 0 & 1 & 0 & 0) \\ (0 & 0 & 0 & 0 & 0 & 0 & 1 & 0) \\ (0 & 0 & 0 & 0 & 0 & 0 & 0 & 1) \end{bmatrix},$$

we have:

$$|H[\eta]\rangle \approx \begin{bmatrix} 2.5018 \\ 1.4511 \\ 1.5508 \\ 0.5788 \\ 0.9877 \\ 0.8941 \\ 0.7219 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

From the above, we have for the finite-L entropy rate convergence:

$$\begin{aligned} h_\mu(L) &= \sum_{N=1}^{\nu_0 - 1} \delta_{L-1,N} \, \langle \delta_\pi | \, W_0 W^N | H(W^{\mathcal{A}}) \rangle \\ &\quad + \sum_{\lambda \in \Lambda_W} \lambda^{L-1} \, \langle \delta_\pi | \, W_\lambda | H(W^{\mathcal{A}}) \rangle \\ &= \sum_{N=0}^{2} \delta_{L-1,N} \overbrace{\langle \delta_\pi | \, W_0}^{= \langle \delta_\pi | - \langle \pi_W |} \, W^N | H(W^{\mathcal{A}}) \rangle \\ &\quad + \underbrace{\langle \delta_\pi | \, W_1}_{= \langle \pi_W |} | H(W^{\mathcal{A}}) \rangle \\ &= h_\mu + \sum_{N=0}^{2} \delta_{L-1,N} \left(\langle \delta_\pi | \, W^N | H(W^{\mathcal{A}}) \rangle - h_\mu \right) \\ &= \delta_{L,1} \, \langle \delta_\pi | \, H(W^{\mathcal{A}}) \rangle + \delta_{L,2} \, \langle \delta_\pi | \, W | H(W^{\mathcal{A}}) \rangle \\ &\quad + \delta_{L,3} \, \langle \delta_\pi | \, W^2 | H(W^{\mathcal{A}}) \rangle + u_{L-4} h_\mu \\ &\approx 0.999 \delta_{L,1} + 0.782 \delta_{L,2} + 0.720 \delta_{L,3} + 0.599 u_{L-4} \ , \end{aligned}$$

where u_{L-4} is the unit step sequence that is zero for $L < 4$ and unity for $L \geq 4$.

For the excess entropy, we find:

$$\begin{aligned} \mathbf{E} &= \sum_{N=1}^{\nu_0 - 1} \langle \delta_\pi | \, W_0 W^N | H(W^{\mathcal{A}}) \rangle \\ &\quad + \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{1-\lambda} \, \langle \delta_\pi | \, W_\lambda | H(W^{\mathcal{A}}) \rangle \\ &= \sum_{N=0}^{2} \overbrace{\langle \delta_\pi | \, W_0}^{= \langle \delta_\pi | - \langle \pi_W |} \, W^N | H(W^{\mathcal{A}}) \rangle \end{aligned}$$

$$= -3 \langle \pi_W | H(W^{\mathcal{A}}) \rangle + \sum_{N=0}^{2} \langle \delta_\pi | W^N | H(W^{\mathcal{A}}) \rangle$$
$$= \langle \delta_\pi | (I + W + W^2) | H(W^{\mathcal{A}}) \rangle - 3h_\mu$$

$$\approx 0.70430 \ .$$

For the transient information, we have:

$$\mathbf{T} = \sum_{N=1}^{\nu_0 - 1} (N+1) \langle \delta_\pi | W_0 W^N | H(W^{\mathcal{A}}) \rangle + \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{(1-\lambda)^2} \langle \delta_\pi | W_\lambda | H(W^{\mathcal{A}}) \rangle$$
$$= \sum_{N=0}^{2} \overbrace{\langle \delta_\pi | W_0}^{= \langle \delta_\pi | - \langle \pi_W |} (N+1) W^N | H(W^{\mathcal{A}}) \rangle$$
$$= -6 \langle \pi_W | H(W^{\mathcal{A}}) \rangle + \sum_{N=0}^{2} \langle \delta_\pi | (N+1) W^N | H(W^{\mathcal{A}}) \rangle$$
$$= \langle \delta_\pi | (I + 2W + 3W^2) | H(W^{\mathcal{A}}) \rangle - 6h_\mu$$

$$\approx 1.12982 \ .$$

The synchronization information is:

$$\mathbf{S} = \sum_{N=1}^{\nu_0 - 1} \langle \delta_\pi | W_0 W^N | H[\eta] \rangle + \sum_{\substack{\lambda \in \Lambda_W \\ |\lambda| < 1}} \frac{1}{1-\lambda} \langle \delta_\pi | W_\lambda | H[\eta] \rangle$$
$$= \sum_{N=0}^{2} \overbrace{\langle \delta_\pi | W_0}^{= \langle \delta_\pi | - \langle \pi_W |} W^N | H[\eta] \rangle$$
$$= -3 \langle \pi_W | H[\eta] \rangle + \sum_{N=0}^{2} \langle \delta_\pi | W^N | H[\eta] \rangle$$
$$= \langle \delta_\pi | (I + W + W^2) | H[\eta] \rangle$$

$$\approx 4.72481 \ .$$

Collecting our results, the scalar complexity measures are:

$$h_\mu = 0.59916 \text{ bits per step},$$
$$C_\mu = 2.50179 \text{ bits},$$
$$\mathbf{E} = 0.70430 \text{ bits},$$
$$\mathbf{T} = 1.12982 \text{ bits-symbols, and}$$
$$\mathbf{S} = 4.72481 \text{ bits}.$$

REFERENCES

[A1] R. G. James, K. Burke, and J. P. Crutchfield. Chaos forgets and remembers: Measuring information creation, destruction, and storage. *Phys. Lett. A*, 378:2124–2127, 2014.

[A2] J. E. Hanson and J. P. Crutchfield. *Physica D*, 103:169–189, 1997.

[A3] D. P. Varn, G. S. Canright, and J. P. Crutchfield. *Phys. Rev. B*, 66(17):174110–3, 2002.