Algebraic Theory of Patterns as Generalized Symmetries

Adam Rupe[1,*] and James P. Crutchfield[2,†]

[1] *Center For Nonlinear Studies and Computational Earth Science, Los Alamos National Laboratory*
[2] *Complexity Sciences Center, Department of Physics and Astronomy, University of California Davis*

(Dated: June 28, 2022)

We generalize the exact predictive regularity of symmetry groups to give an algebraic theory of patterns, building from a core principle of future equivalence. For topological patterns in fully-discrete one-dimensional systems, future equivalence uniquely specifies a minimal semiautomaton. We demonstrate how the latter and its semigroup algebra generalizes translation symmetry to partial and hidden symmetries. This generalization is not as straightforward as previously considered. Here, though, we clarify the underlying challenges. A stochastic form of future equivalence, known as predictive equivalence, captures distinct statistical patterns supported on topological patterns. Finally, we show how local versions of future equivalence can be used to capture patterns in spacetime. As common when moving to higher dimensions, there is not a unique local approach, and we detail two local representations that capture different aspects of spacetime patterns. A previously-developed local spacetime variant of future equivalence captures patterns as generalized symmetries in higher dimensions, but we show this representation is not a faithful generator of its spacetime patterns. This motivates us to introduce a local representation that is a faithful generator, but we demonstrate that it no longer captures generalized spacetime symmetries. Taken altogether, building on future equivalence, the theory defines and quantifies patterns present in a wide range of classical field theories.

Keywords: Organization, Structure, Pattern, Computational mechanics, Spacetime, Translation symmetry, Predictive equivalence

I. PATTERNS IN NATURE

Symmetry plays a central role in fundamental physics. When we look out at the world around us, on the human scale, however, there is a notable lack of exact symmetries. Cows are not spherical, for instance. The disconnect between physics at the fundamental level and the human scale is often described in terms of *spontaneous symmetry breaking*. How and why spontaneous symmetry breaking occurs so ubiquitously in natural systems are interesting and challenging questions, but not our concern here. Rather, we are interested in the question of what structures result from broken symmetries. In particular, a special case of spontaneous symmetry breaking is *spontaneous self-organization*. But what *is* organization in the first place? Can we mathematically formalize it and quantitatively measure it?

We will use the general rubric *pattern* to refer to the forms of organization that spontaneously emerge in natural systems. When a system undergoes a spontaneous symmetry breaking event it self-organizes into some pattern, either in time, space, or both.

In many canonical examples of *pattern formation* near equilibrium [1–3]—recall the convection roles that emerge in Bénard flow [4, 5] or the spiral waves in the Belousov-Zhabotinsky chemical reaction [6, 7]—a system undergoes a continuous-to-discrete symmetry breaking bifurcation event. This occurs when it self-organizes, going from a homogeneous state with trivial continuous spacetime symmetries to a state with nontrivial discrete symmetries.

Hexagonal convection cells form in a fluid with velocity initially zero everywhere during a Bénard instability. Belousov first discovered a "chemical clock" with a discrete-time symmetry oscillation that arises from an initially stationary mixture. Farther from equilibrium, these discrete symmetries may be further broken during subsequent bifurcations, resulting in states we may consider "patterned", but that have no discernible simple symmetries remaining. Turbulent fluid flows containing coherent structures [8, 9], such as Jupiter's Great Red spot [10, 11], provide many commonly-encountered examples.

What we would like then, and what the following contributes, is a unified account of patterns—an account that rigorously and formally describes the full range of phenomenon including, but going beyond, organizations with exact symmetries. Given that symmetries are formally captured using the mathematics of group theory and given the enormous success group theory has had in formulating physical theory [12], we take an algebraic approach to framing patterns as generalized symmetries.

We start with the simplest setting of discrete one-dimensional spatial systems (e.g., spin lattices) and show how the semigroup algebra of semiautomata—a mathematical representation originating in symbolic computation—generalizes translational symmetry. In doing so, we rigorously clarify subtleties posed by this generalization—subtleties that have not been previously addressed. Based on this, we introduce a classification hierarchy in terms of exact symmetries, partial symme-

* adamrupe@lanl.gov
† chaos@ucdavis.edu

tries, hidden symmetries, and general patterns. We also describe distinct statistical structures supported on these one-dimensional patterns and show that stochastic generalizations of semiautomata provide mathematical representations of these statistical structures.

In addition, we explore generalizations to patterns in higher dimensions. We introduce a class of local models that generalize the semiautomata approach for spatiotemporal systems. Two models in this class are shown to be particularly useful. Intriguingly, the utility of these models are apparently mutually exclusive. The first model, introduced previously, can discover hidden spacetime symmetries and coherent structures [13], such as those in turbulent fluid flows [14]. However, we describe a previously unknown shortcoming of these models: they are not consistent generators of their associated spacetime field patterns. The second model, introduced here for the first time, corrects this flaw and introduces a consistent generator of spacetime field patterns. Unfortunately, it loses the first model's useful generalized spacetime symmetries.

These representations' conflicting strengths—generalized spacetime symmetries and consistent spacetime generation—add new questions and suggest new paths of inquiry to the enigma of patterns in higher dimensions [15–17].

II. ONE-DIMENSIONAL PATTERNS

Abstractly, we can think of a pattern as a *predictive* regularity [18]:

> ... some object O has a pattern P—O has a pattern 'represented', 'described', 'captured', and so on by P—if and only if we can use P to predict or compress O.

On one extreme, symmetries represent an *exact* predictive regularity. If the symmetries are known, the pattern can be perfectly predicted at any other point in time or space. On the other extreme, a completely random system is entirely devoid of predictive regularity. If every point in spacetime is an independent, identically distributed random variable, there is no regularity. And so, knowledge of any part of the system cannot be leveraged to predict other parts of the system. The notion of pattern that we seek encompasses both of these extremes and systems in between. A general pattern will be neither perfectly predicable nor entirely unpredictable—it will be an amalgamation of regularity and randomness.

Before proceeding, let us briefly compare and contrast the theory developed here with the Pattern Theory of Ulf Grenander and colleagues [19]. As both aim at a general quantitative understanding of what patterns are and how to discover them in the world, there are many conceptual similarities. While some quantitative similarities emerge, in particular the use of nonparametric learning algorithms for hidden Markov models [20], most of the quantitative machinery differs. Pattern Theory is grander in scope than what is developed here, and so we do not require the very general constructs of *bonds*, *connectors*, *configurations*, *images*, and the like [19]. Likewise, Pattern Theory does not employ the machinery of symbolic dynamics, sofic shifts, and predictive equivalence used here. Thus, our work is complementary to the more general approach of Pattern Theory.

A. Statistical Field Theories

The following mainly concerns fully-discrete one-dimensional spatial systems. These are given as a *shift space* $\mathcal{X}$—a set of indexed bi-infinite sequences, or *strings*, of symbols taken from a finite alphabet $\mathcal{A}$. Before diving into details, let's first take a moment to compare shift spaces to the analogous setup from statistical mechanics for analyzing ordered systems.

A shift space can be thought of as a *topological ensemble*—a set of strings—in contrast with a statistical ensemble that is a distribution over a set of strings. This is an abstraction of discrete-spin models in one-dimension—e.g., $\mathcal{A} = \{-1, 1\}$ for a standard Ising model. Rather than specify interactions on the spin lattice and analyze the resulting statistical field theory, we wish to analyze any pattern present for a given (topological) ensemble $\mathcal{X}$.

A key distinction between a shift space $\mathcal{X}$ as a topological ensemble and a spin lattice ensemble in a statistical field theory is that all elements $x \in \mathcal{X}$ are related to one another through the shift operator σ. (Formally described below.) In fact, for the irreducible sofic shifts we consider, $(\mathcal{X}, \sigma)$ is an ergodic dynamical system. And so, every member x of the ensemble can eventually be sampled through σ's action. Thus, we equivalently consider (i) $\mathcal{X}$ as an ensemble of points and σ as a deterministic mapping between those points or (ii) $\mathcal{X}$ as a single infinite lattice and σ moves indices on that lattice. The difference is that of active versus passive transformation.

Spontaneous symmetry breaking in statistical field theories is monitored through an *order parameter*; such as total magnetization for an Ising model. In the symmetric "ordered" phase, the order parameter has a nonzero value and, after a symmetry-breaking phase transition, the order parameter vanishes. For the Ising model, below the transition—below the *critical temperature*—spins tend to align giving nonzero magnetization. At zero temperature the model reaches its ground state with all spins aligned. This is a fully symmetric state with maximal magnetization, corresponding to strings of the form $\{\ldots, 1, 1, 1, \ldots\}$ or $\{\ldots, -1, -1, -1, \ldots\}$. Above the critical temperature, this symmetry is fully broken, with zero magnetization.

While effective as an approach to thermal spin lattices, such as the Ising and related lattice models, abstractly quantifying "order" with a single scalar quantity—the order parameter—is far from ideal.

First, for the Ising model, there are only two configura-

tions with maximal magnetization, as given above. Second, these configurations are maximally symmetric, with $\sigma^p(x) = x$ for integer p. Consider, though, configurations of the form $\{\ldots, -1, 1, -1, 1, -1, \ldots\}$. These configurations are still symmetric, with $\sigma^{2p}(x) = x$, although they have vanishing order parameter. There are many such symmetric configurations with zero magnetization: e.g., those of the form $\{\ldots, (-1)^n, (1)^n, \ldots\}$, with $\sigma^{2np}(x) = x$. More novelly, there are zero order-parameter configurations that are neither completely symmetric nor completely random. Third, these symmetric sequences with zero order parameter are not the ground state and they are not stable under thermal perturbations. Thus, though singled out by the choice of total magnetization as *the* order parameter, they are edge cases that will almost never be seen. Finally and more generally, order parameters in statistical mechanics are not determined from first physical principles. They must be posited initially and then proved appropriate.

Similarly, *correlation functions* and *structure factors* are additional and commonly-employed scalar quantities that capture one or another notion of order. Conceptually, a system considered highly ordered will surely be highly correlated. Patterns that emerge on the macroscopic scale correspond to collective behaviors on the microscale that certainly exhibit nonzero correlations. However, as with order parameters, there can be many degeneracies between specific patterns and correlation values. In short, order is something beyond correlation. A diverging correlation length in an Ising model at the critical temperature does not signify the presence of intricate patterns and organization, such as spiral wave patterns in lattice models of excitable media [21]. To remedy these failings we seek a definition of pattern that is not scalar.

This is not to banish all scalar quantities. Many, in given settings, can be insightful [22–25]. We will show that the algebraic *presentations* for topological patterns have a natural extension to patterns in statistical field theories. Moreover, scalar quantities of interest, like correlation functions, can be computed in closed-form from the stochastic presentations. We also return later, briefly, to discuss generalized order parameters in light of the algebraic theory.

B. Symbolic Dynamics

We now detail shift spaces and how they quantify topological patterns as generalized symmetries. Consider a finite alphabet of n symbols $\mathcal{A} = \{0, 1, \ldots, (n-1)\}$ and (indexed) bi-infinite symbol sequences or *strings*. The set $\mathcal{A}^{\mathbb{Z}}$ of all possible bi-infinite sequences is known as the *full-n shift*. A particular sequence $x = \ldots x_{-1}x_0x_1 \ldots \in \mathcal{A}^{\mathbb{Z}}$ is described as a *point* in $\mathcal{A}^{\mathbb{Z}}$. For now, we need not specify whether sequence indices are time coordinates or space coordinates. In either case, translations are generated via the *shift operator* σ that maps a point $x \in \mathcal{A}^{\mathbb{Z}}$ to another point $y = \sigma(x)$ whose i^{th} coordinate is $y_i = x_{i+1}$ for all i. (That is, σ shifts every element of x one place to the left.) Our interest is in patterns as predictive regularity, and the predictions are made over translations generated by σ. Thus, we want to capture patterns in closed, σ-invariant subsets of $\mathcal{A}^{\mathbb{Z}}$. The subsets are called *shift spaces* (or *subshifts* or simply *shifts*).

Often one can concisely specify a shift space as the set of all strings that do not contain a collection of *forbidden words*. A *word* is a finite block of symbols $a_j \in \mathcal{A}$ and a point x is said to contain or admit a word $w = a_0a_1 \cdots a_k$ if there are indices i and $j = i + k$ such that $x_{i:j} = w$; explicitly, $x_i = a_0, x_{i+1} = a_1, \ldots, x_{j-1} = a_k$. Again, a word is a finite sequence of symbols; a string, bi-infinite.

For a collection $\mathcal{F}$ of forbidden words, define $\mathsf{X}_{\mathcal{F}}$ to be the subset of strings in $\mathcal{A}^{\mathbb{Z}}$ that do *not* contain any words $w \in \mathcal{F}$. A shift space $\mathcal{X}$ is a subset of the full shift $\mathcal{A}^{\mathbb{Z}}$ such that $\mathcal{X} = \mathsf{X}_{\mathcal{F}}$ for some collection $\mathcal{F}$ of forbidden words [26]. The *language* $\mathcal{W}(\mathcal{X}) = \mathcal{F}^c$ of a shift space $\mathcal{X}$ is the collection of all words that occur in some point in $\mathcal{X}$.

If $\mathcal{F}$ is a finite set, the resulting shift space is called a *subshift of finite type* [27] or an *intrinsic* or a *topological Markov chain* [28]. A wider class of finitely-definable shift spaces are the *sofic shifts*. These are the closure of subshifts of finite type under continuous local mappings—*k-block factor maps* [29]. Though, note that there are many equivalent definitions of sofic shifts; several of which are given below as needed. A sofic shift is *irreducible* if, for every ordered pair of words $u, v \in \mathcal{W}(\mathcal{X})$, there is a word w such that $uwv \notin \mathcal{F}$. For reasons elaborated on shortly, we define general discrete one-dimensional patterns as irreducible sofic shifts.

C. Sofic Shifts as Topological Patterns

To recap, we seek a mathematical specification of patterns in strings that captures a range of organizations spanning fully-symmetric sequences to arbitrary ("random") sequences. Moreover, we wish to identify, from first principles, an associated algebra that generalizes the group algebra of symmetries. For physical consistency, we started with shift spaces since they are shift-invariant subspaces of strings. To fulfill the algebraic requirement, we now further restrict to sofic shifts as they are shift spaces defined in terms of a finite semigroup [29].

Recall that a *group* is a set of elements closed under an associative and invertible binary operation with an identity element. In this, they are too restrictive and impose only exact symmetry. In contrast, *semigroups* require neither invertibility nor an identity operation. This relaxation is key to defining generalized patterns, as we laid out above. It permits exact symmetries but also allows expressing noisy and approximate symmetries.

The elements of a sofic shift's semigroup are words and the binary operation is word concatenation. For example, the set $\mathcal{W}(\mathcal{A}^{\mathbb{Z}})$ of all words in $\mathcal{A}^{\mathbb{Z}}$ and their concatenation products together form the *free semigroup*. For

example, the product of $u = 00$ and $v = 11$ in the free semigroup is the word $w = uv = 0011$. A sofic shift $\mathcal{X} = \mathsf{X}_{\mathcal{F}}$ is defined in terms of a finite semigroup G with an absorbing element e, whose product is $ge = eg = e$ for all $g \in G$. The absorbing element e together with the elements from the alphabet $\mathcal{A} = \{0, 1, \ldots, (n-1)\}$ generates G via single-symbol concatenations. G's production rules are such that for any pair of allowed words $u, v \notin \mathcal{F}$, if their concatenation $w = uv$ contains a forbidden word $f \in \mathcal{F}$, then their product in G gives the absorbing element $uv = e$.

A semigroup of word concatenations is also associated with a simple presentation in the form of a *semiautomaton* finite-state machine [30]—the triple $(\Xi, \mathcal{A}, \mathcal{M})$, where $\Xi = \{\xi_0, \xi_1, \ldots, \xi_m\}$ is a set of *internal states*, $\mathcal{A} = \{0, 1, \ldots, (n-1)\}$ is the symbol alphabet, and $\mathcal{M} = \{M_0, M_1, \ldots, M_{(n-1)}\}$ is a set of mappings from Ξ into Ξ. To be explicit, we consider *deterministic and fully-specified semiautomata* for which each M_a is a function—each input has one and only one output—with domain over the full set Ξ of internal states. Semiautomata can be usefully depicted as an edge-labeled directed graph. The vertices represent the internal states in Ξ and for every pair (ξ_i, ξ_j) such that $\xi_j = M_a(\xi_i)$ there is an edge labeled $a \in \mathcal{A}$ that leads from ξ_i to ξ_j. For a deterministic and fully-specified semiautomaton, there is an edge labeled with each symbol in $\mathcal{A}$ emanating from every state in Ξ.

A fully-specified semiautomaton directly determines a subshift's algebra from the free semigroup as follows. For every state $\xi_i \in \Xi$ and any element $a_0 a_1 \cdots a_k$ of the the free semigroup $\mathcal{A}^{\mathbb{Z}}$ there is a map $M_{a_0} \circ M_{a_1} \circ \cdots \circ M_{a_k}$ from ξ_i to another state $\xi_j \in \Xi$ [30]. A deterministic and fully-specified semiautomaton is a *presentation* of a sofic shift if we include an absorbing "forbidden" state $\xi_e \in \Xi$. That is, the mappings associated with all elements of the free semigroup containing a forbidden word in $\mathcal{F}$ lead to the forbidden state. And, all mappings from the forbidden state return the forbidden state. That is, $M_w(\xi_i) = \xi_e$ for all ξ_i and $w \in \mathcal{F}$, and $M_a(\xi_e) = \xi_e$ for all $a \in \mathcal{A}$. See Fig. 1 for presentations of example shifts.

Since sofic shifts are defined by a finite semigroup, every sofic shift can be presented by a semiautomaton with a finite set of states Ξ. Recall from above that the idea of *compression* is related to our intuition of pattern. A pattern—a predictive regularity—allows for a compressed representation of a system's behaviors. Note that sofic shifts and their presentations provide a finite representation of an ensemble of infinite strings through their finite semigroup.

Here, we distinguish between three types of sofic-shift semiautomaton presentation. Appendix A gives example constructions of these three types of presentation.

The most straightforward presentation assigns a state $\xi \in \Xi$ to each element of a semigroup G of $\mathcal{X}$ and fills in the state transitions M_a using G's production rules [31]. While straightforward to construct, if a G is known, this semiautomaton presentation is not necessarily minimal, in terms of the state set size $|\Xi|$. To specify a particular pattern as a sofic shift, it is crucial to have a minimal and unique presentation associated with the given sofic shift. This also allows extracting unambiguous quantitative measures of the ensemble of strings, such as measures of correlation, from the finite presentation.

An important presentation that is minimal and constructible without knowing any G is $\mathcal{X}$'s *future cover* [32], defined below. The future cover semiautomaton of every irreducible sofic shift $\mathcal{X}$ has a unique strongly connected component [33, 34]. This irreducible component is our mathematical representation of patterns as generalized symmetries. We refer to it as the *canonical machine presentation* $\mathcal{P}(\mathcal{X})$ *of* $\mathcal{X}$. The future cover and its recurrent component $\mathcal{P}(\mathcal{X})$ provide a unique, minimal mathematical representation of $\mathcal{X}$.

The *future set* (sometimes *follower set*) $F_{\mathcal{X}}(w)$ of a word $w \in \mathcal{W}(\mathcal{X})$ is the collection of all words u such that $wu \in \mathcal{W}(\mathcal{X})$. Define the *future equivalence relation* $\sim_F$ on $\mathcal{X}$ as:

$$u \sim_F v \iff F_{\mathcal{X}}(u) = F_{\mathcal{X}}(v) \ , \tag{1}$$

for all $u, v \in \mathcal{W}(\mathcal{X})$.

An important definition of sofic shifts that we use shortly is:

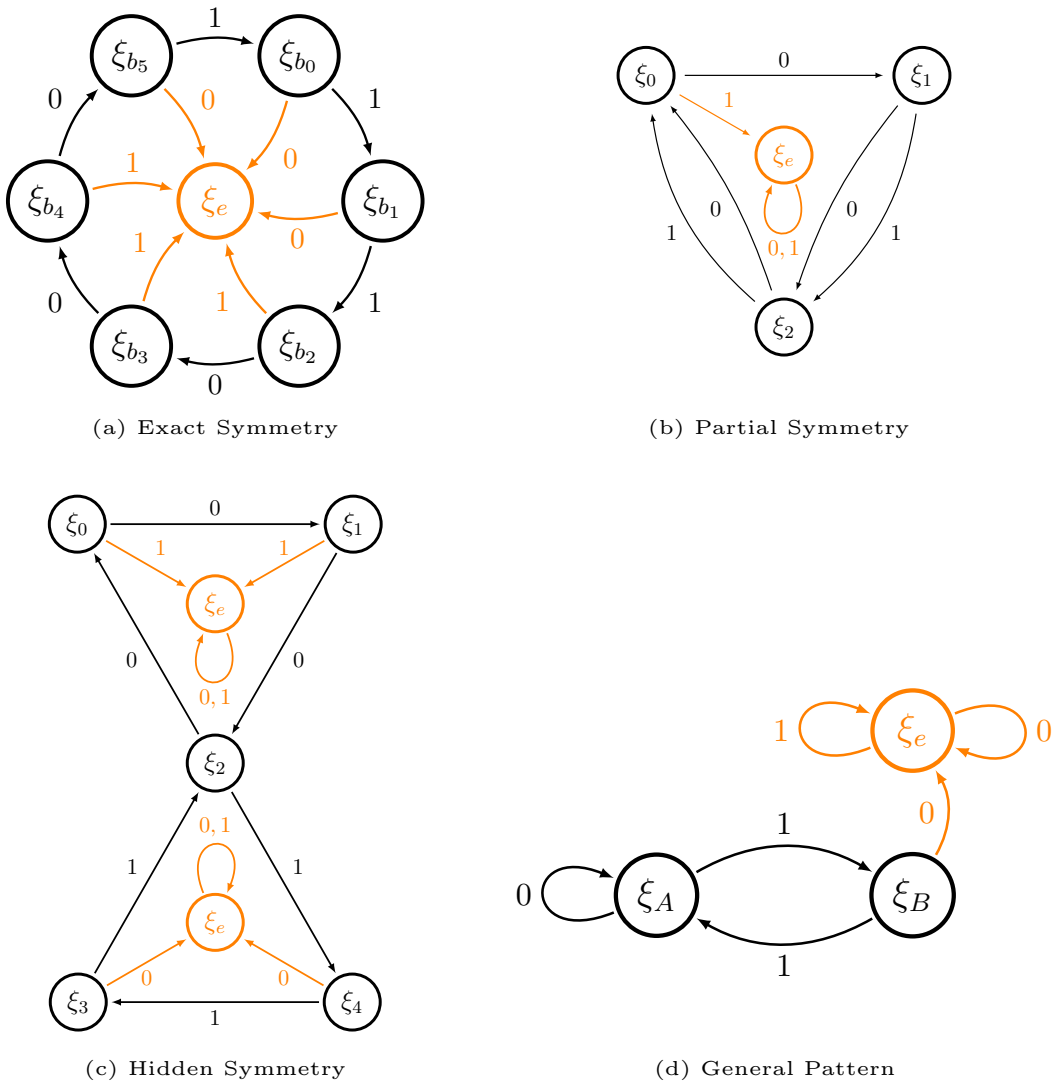
Theorem (1)**.** *[26, Theorem 3.2.10] A shift space is sofic if and only if it has a finite number of future sets.*

Therefore, sofic shifts have finitely-many equivalence classes, denoted $[\cdot]_F$, and these equivalence classes plus the absorbing forbidden state are the internal states Ξ of the future cover semiautomaton.

The mappings M_a that give the state transitions are defined from the allowed concatenations that do not contain a forbidden word in $\mathcal{F}$. That is, each state $\xi_i \in \Xi \setminus \{\xi_e\}$ is an equivalence class $[u]_F$ and, for each symbol $a \in \mathcal{A}$, the concatenation $v = ua \neq e$ belongs to the equivalence class $[v]_F$ assigned as state $\xi_j \in \Xi$, giving $M_a(\xi_i) = \xi_j$. Note that this is independent of the choice $u \in [u]_F$ and that $[v]_F$ in some cases may be equal to $[u]_F$. This gives a self-edge transition in the semiautomaton: $M_a(\xi_i) = \xi_i$. If $ua = e$, then $M_a(\xi_i)$ maps to the forbidden state ξ_e.

This natural transition structure follows from future equivalence and it leads to a very important property of the future-cover semiautomata. They are called *unifilar* [35] in information theory, equivalently also called *right resolving* in symbolic dynamics [26, Corollary 3.3.19] or *deterministic* in automata theory [36]. A fully-specified semiautomata is unifilar if for every internal state ξ_i and every word (element of the free semigroup) $a_0 a_1 \cdots a_k$ the map $M_{a_0} \circ M_{a_1} \circ \cdots M_{a_k}$ leads to one and only one internal state ξ_j. (It may be that $j = i$.)

Unifilarity is the defining property of a *predictive* semiautomaton. Since the goal is to formalize patterns as predictive regularities, this is an important point to stress. By way of contrast, first note that any presentation of



(a) Exact Symmetry

(b) Partial Symmetry

(c) Hidden Symmetry

(d) General Pattern

FIG. 1. Semiautomata presentations $\mathcal{P}(\mathcal{X})$: (a) Exact symmetry shift, (b) partial symmetry shift, (c) hidden symmetry shift, and (d) general pattern sofic shift.

a sofic shift $\mathcal{X}$, whether unifilar or not, is a *generator* of $\mathcal{X}$. Every string in $\mathcal{X}$ can be generated by following the symbol-labeled transitions of the presentation and no forbidden words can be generated. Thus, being generative can be thought of as the baseline property of any model of a sofic shift. Prediction is an additional capability beyond generation that arises from a presentation being unifilar.

Unifilarity establishes a presentation as predictive in the following way. For topological patterns, the task of prediction is to establish what may happen in the future, given what has happened in the past. Specifically, given a word w, what words are allowed to follow in the shift space? That is, what is w's future set? Due to the implied determinism, each internal state of a unifilar presentation is uniquely specified by the word w leading to that state. Notably, this is not guaranteed for an arbitrary generator of a shift. Furthermore, in a unifilar presentation every subsequent word again uniquely leads to another internal state. It is straightforward to see that the future cover is a predictive presentation: By definition, its internal states are *future separated*—the set of all words that may follow from each internal state is unique.

To establish that sofic shifts and their canonical machine presentations express patterns as generalized symmetries, it is helpful to first describe how they capture exact translation symmetries of symbolic sequences.

D. Exact Symmetries

A string x has a discrete translation symmetry if $\sigma^p(x) = x$, where the minimal such $p \in \mathbb{N}$ is the symmetry's *period*. The symmetry group is the set $\{\sigma^{np} :$

$n \in \mathbb{N}\}$ with $\sigma^{ip}\sigma^{jp} = \sigma^{kp}$ where $k = i + j$. Since here p is finite, the action of σ on x produces a compact shift-invariant subspace of $\mathcal{A}^{\mathbb{Z}}$. Therefore, translation symmetric strings are shift spaces $\mathcal{X}$.

We now show that the shift space $\mathcal{X}$ of translation symmetric strings is sofic. The action of the symmetry group used to define x is determined by the shift operator σ, while the action of sofic semigroups is word concatenation.

To connect these, consider *windows* $x_{i:j}$ $(i < j)$ that return the word w from coordinates i through j in x. For $(i < j < k)$, if $x_{i:j} = u$ and $x_{j:k} = v$, then $x_{i:k} = w = uv$ gives the concatenated window. Recall that σ shifts indices in x so that $\sigma^k(x)_i = x_{i+k}$. If we have a word $u \in \mathcal{W}(\mathcal{X})$, there is some (i, j) such that $x_{i:j} = u$. Then the allowed concatenations $uv \neq e$ are determined by the shift operator σ, since we can write $v = x_{j:j+k} = (x_j)\sigma(x)_j\sigma^2(x)_j \cdots \sigma^k(x)_j$. All such concatenations determine $\mathcal{X}$'s semigroup G. For translation symmetric strings with $\sigma^p(x) = x$, $\sigma^p(x)_i = x_i$. And so, intuitively, there is only a finite number of elements in G. Therefore, $\mathcal{X}$ for translation symmetric strings is sofic.

We now show this explicitly by constructing the canonical machine presentation $\mathcal{P}(\mathcal{X})$ with a finite number of states.

Proposition. *Translation-symmetric strings, $\sigma^p(x) = x$ for some $p \in \mathbb{N}$, together with their shifts $y = \sigma^n(x)$ for all $n \in \mathbb{N}$, form an irreducible sofic shift space.*

Proof. First, note that $\sigma^p(x) = x$ implies x can be written as a tiling $\cdots bbbbb \cdots$, where b is a word of length p. Pick any x_i as the first symbol b_0 in the word b. Then $\sigma(x)_i = b_1$, $\sigma^2(x)_i = b_2$, ..., with $\sigma^{p-1} = b_p$. Applying σ one more time gives $\sigma^p(x)_i = x_i = b_0$, arriving at the next tile b.

Second, using this observation we create $\mathcal{P}(\mathcal{X})$ using $p + 1$ internal states, where we have a state ξ_{b_i} for each symbol $b_i \in \mathcal{A}$ in the tile b (there are p of these) and one absorbing state ξ_e. Let ξ_{b_0} be the future set equivalence class $[b]_F$ of the word b. Due to x's exact symmetry, there is one and only one symbol $a \in \mathcal{A}$ such that $ua \neq e$ for $u \in [b]_F$; namely b_1. Therefore, let $M_{b_1}(\xi_{b_0}) = \xi_{b_1}$, and then all other M_a map to ξ_e, for $a \neq b_1$. Now, $\xi_{b_1} = [bb_1]_F$ and b_2 is the only symbol such that $ua \neq e$ for $u \in [bb_1]_F$. Thus, $M_{b_1}(\xi_{b_1}) = \xi_{b_2}$ and all other M_a map to ξ_e, for $a \neq b_2$. Repeat this argument for all generators in b, with $\xi_{b_p} = [bb_1b_2 \ldots b_p]_F = [bb]_F$.

Third, as with ξ_{b_0}, the only symbol that can follow is b_1 and we repeat the full argument again, where each b_i sequentially follows. Therefore, $[bb]_F$'s future set equals that of $[b]_F$ and so $[bb]_F = [b]_F$. Thus, the only transition from ξ_{b_p} that is allowed returns to the original state ξ_{b_0}. This completely specifies $\mathcal{P}(\mathcal{X})$ with a finite number of states $\Xi = \{\xi_{b_0}, \ldots, \xi_{b_p}, \xi_e\}$. Given Theorem (1) above, one concludes that the shift space $\mathcal{X}$ for a translation symmetric sequence is sofic. □

Figure 1(a) gives an example of $\mathcal{P}(\mathcal{X})$ for the set of translation symmetric strings $\cdots 000111000111000 \cdots$ with $b = 111000$. For visual clarity it omits the self-loops on state ξ_e.

Recall that $\mathcal{P}(\mathcal{X})$ is the irreducible component of the future cover, so for a given shift there may be additional states beyond those just described. However, these are equivalence classes for words of length less than p, since we started with $[b]_F$. Since p is finite, there are only finitely many additional equivalence classes. These correspond to transient (nonrecurrent) states of the future cover semiautomaton.

In Fig. 1(a)'s example, knowing $x_i = 0$ does not fully specify which state $\mathcal{P}(\mathcal{X})$ is in, since three states ξ_3, ξ_4, and ξ_5 correspond to observing symbol 0. The additional transient states of the future cover specify how to *synchronize* to $\mathcal{P}(\mathcal{X})$'s states from the generators $a \in \mathcal{A}$ (single-symbol words).

Having constructed the canonical machine presentation $\mathcal{P}(\mathcal{X})$, we can further relate $\mathcal{X}$'s semigroup action to x's translation symmetry group. For each state $\xi_{b_i} \in \Xi$ there is one and only one transition that does not lead to the absorbing forbidden state ξ_e. That is, only one generator $a \in \mathcal{A}$ can be concatenated to the words $u \in [u]_F = \xi_{b_i}$. Similarly, if we consider a word $u \in \xi_{bi}$ as the window $u = x_{i:j}$, then a unit shift by σ reveals one and only one new symbol at index j in $\sigma(x)_{i:j}$. Therefore, ignoring the absorbing state and transitions to it, the graph of $\mathcal{P}(\mathcal{X})$ is a cyclic graph with period p: Every p-length path from ξ_i returns to ξ_i for all $\xi_i \in \Xi \setminus \{\xi_e\}$. Thus, the permutation symmetries of this (edge-labeled) graph correspond to elements of x's translation symmetry group.

In Fig. 1(a)'s example, the state labeled ξ_{b_0} corresponds to the start of the tile $b = 111000$, but we could equivalently use ξ_{b_0} as the start of tile 000111. Furthermore, the internal states have a functional meaning. They are the elements of the quotient group of the translation symmetry—counters that track the symmetry's phase.

It must be emphasized that sofic shifts and their semigroups do not formally generalize such exact symmetries in the obvious way. That is, the semigroup of a sofic shift for exact symmetry strings does not become a symmetry group. G's absorbing semigroup element e is still required for an exact symmetry sofic shift $\mathcal{X}$. From our construction of translation symmetric $\mathcal{P}(\mathcal{X})$ we see that, for every internal state $\xi_{b_i} \neq \xi_e$, there is one and only one transition that does not lead to ξ_e. This makes it clear that exact symmetries are a highly restrictive form of pattern. By representing exact symmetries using sofic shifts and their machine presentations, though, it is now straightforward to generalize by relaxing the restrictions that impose exact symmetries.

E. Generalized Symmetries

Ignoring ξ_e, a sofic shift $\mathcal{X}$ whose machine presentation is not cyclic then represents a pattern as a generalized

symmetry. We can again either consider $\mathcal{X}$ as an ensemble of strings or one infinite sequence that possesses a generalized symmetry described by $\mathcal{X}$'s semigroup, which is well represented by the machine presentation $\mathcal{P}(\mathcal{X})$. By removing the restriction of a cyclic graph in $\mathcal{P}(\mathcal{X})$—that imposed the perfect regularity of $x = \cdots bbbbb \cdots$—we now can capture a much wider class of patterned strings with approximate or partial regularities.

Consider the extreme case of the full shift $\mathcal{A}^{\mathbb{Z}}$ that has no regularity. There are no products $uv = e$ in G for $\mathcal{A}^{\mathbb{Z}}$, and so there are no restrictions on its words. The algebra of the full shift is the free semigroup. Its machine presentation $\mathcal{P}(\mathcal{A}^{\mathbb{Z}})$ is a single state with all M_a mapping that one state back to itself—i.e., all transitions are self transitions. Since all words can be concatenated to each other, they all belong to a single future equivalence class.

We interpret $\mathcal{A}^{\mathbb{Z}}$'s complete lack of regularity to be a null pattern. Analogously, the opposite extreme of total regularity with strings of the form $x = \cdots aaaaa \cdots$ for $a \in \mathcal{A}$ is also a null pattern with a single-state (again, ignoring ξ_e) machine presentation that has a single self-transition. The null pattern, in these cases, has zero memory—the logarithm of $|\Xi \backslash \{\xi_e\}|$ vanishes. While both extremes at first seem to be polar opposites, recall our goal is that "pattern" represents a *predictive* regularity and this is lacking in both cases. The full shift is completely "random" and thus unpredictable. Whereas, for trivially translation symmetric strings $x = \cdots aaaa \cdots$, the future is always the same. There is nothing to predict.

Between the complete regularity of exact symmetries and lack of predictive regularity in null patterns, we identify several categories of partial predictive regularity.

First, note that for translation symmetric strings $\sigma^p(x)_i = x_i$ for all i.

Second, there are string classes for which $\sigma^p(x)_i = x_i$ for only *some* i. We call these *partial symmetries*. A particular case of partial symmetries are *stochastic symmetries*. For simplicity, consider binary sequences with $\mathcal{A} = \{0, 1\}$, and let ω denote a "wildcard" that can be either 0 or 1 [37]. Sofic shifts with stochastic (partial) symmetries are fully translation symmetric after making wildcard substitutions. For example, we can specify a sofic shift with sequences of the form $x = \cdots \omega\omega\ 0\ \omega\omega\ 0\ \omega\omega\ 0\ \cdots$, say. Examples of such strings are $\cdots 11\ 0\ 01\ 0\ 00\ 0\ 10\ 0\ 11\ 0 \cdots$, where spaces help emphasize the "fixed" 0s that are the scaffolding of the partial symmetry. Note that the canonical machine presentation $\mathcal{P}(\mathcal{X})$ for such stochastic symmetries are also cyclic graphs, as shown in Fig. 1(b).

Third, recall that if the canonical presentation $\mathcal{P}(\mathcal{X})$ is a cyclic graph, every p-length path from ξ_i returns to ξ_i for all $\xi_i \in \Xi \setminus \{\xi_e\}$. Similar to how we generalized from exact to partial symmetries, we define *hidden symmetries* for which only *some* states $\xi_i \in \Xi \setminus \{\xi_e\}$ return to themselves on all p-length paths in $\mathcal{P}(\mathcal{X})$. We exclude the case of $p = 1$, so that self-loops do not count as a hidden symmetry. Figure 1(c) shows an example with ξ_2 as the symmetric state. The canonical machine presentation specifies a sofic shift consisting of arbitrary arrangements of blocks $a = 000$ and $b = 111$; e.g., $x = \cdots aababba \cdots = \cdots 000000111000111111000 \cdots$. The exact symmetry shift in Fig. 1(a) is the special case of the symmetric tiling $x = \cdots abababa \cdots$.

Finally, Fig. 1 (d) gives an example of a general nonnull pattern that is not an exact, partial, or hidden symmetry. This is the well-known Even Shift [29, 31]—the set of binary strings in which only even blocks of 1s bounded by 0s are allowed. This $\ldots 01^{2n}0 \ldots$ pattern extends to arbitrary lengths, despite being specified by only two internal states. While there are no states in the presentation $\mathcal{P}(\mathcal{X})$ that always return to themselves after $p \neq 1$ transitions, there is still predictive regularity. In particular, if a 1 is seen after a 0, it is guaranteed that the next symbol will be a 1. This is specified by ξ_B having only one allowed transition to ξ_A on a 1. Appendix A 2 discusses this example in more detail, along with its semigroup and three semiautomaton presentations.

Before moving to probabilistic patterns represented by measure sofic shifts, we note that Krohn-Rhodes theory [38, 39] was the first to connect finite automata with a semigroup algebra. Moreover, it showed that finite semigroups and their corresponding automata naturally decompose into simpler components, including finite simple groups. This is yet another perspective showing finite automata and their semigroup algebra capture patterns as generalized symmetries. Further exploration of the connection between Krohn-Rhodes theory and the perspective developed here is left for future work. One important difference to note is that their approach did not address statistical or noisy patterns, as the following now does.

F. Statistical Patterns Supported on Sofic Shifts

The exposition on sofic shift patterns did not invoke probabilities over symbols, words, or strings. Shifts spaces are not concerned with the probability of a word occurring, only whether a word can possibly occur or not. This is why we referred to sofic shifts as *topological* patterns. We just saw that exact symmetries are given as sofic shifts and so are topological patterns. Recall that our key motivation for showing this was to argue for sofic shifts as a mathematical formalism that captures a notion of (topological) pattern that greatly expands exact symmetry to generalized symmetry. However, as we now describe, we can generalize further to formalize *statistical* patterns that are supported on sofic shifts. Doing so provides a direct link with the more familiar statistical measures of order and organization used in statistical mechanics, such as correlation functions.

Our goal is to build a probability space on top of a shift space. The key property of this probability space is that words are assigned positive probability if and only if they are allowed in the shift space. This is accomplished through the use of cylinder sets as the sigma algebra on

a shift space $\mathcal{X}$ [40]. For a shift space $\mathcal{X}$ and a length-n word ω, the *cylinder set* $C_i(\omega)$ is defined as:

$$C_i(\omega) := \{x \in \mathcal{X} \ : \ x[i : i+n-1] = \omega\} \ .$$

Naturally then, a probability measure μ is assigned such that the probability of a word ω occurring is given as:

$$\Pr(\omega) = \mu\big(C_i(\omega)\big) \ .$$

By definition, $\Pr(\omega) = 0$ if $\omega \in \mathcal{F}$, as the associated cylinder sets will be empty for forbidden words.

Such a probability measure defines a *stationary stochastic process* over the shift space $\mathcal{X}$ if it is shift-invariant, such that the probability of a word is independent of the index i in $C_i(\omega)$, and each word satisfies prefix and suffix marginalization:

$$\mu(w) = \sum_{\{a:aw \in \mathcal{W}(\mathcal{X})\}} \mu(aw)$$

and:

$$\mu(w) = \sum_{\{a:wa \in \mathcal{W}(\mathcal{X})\}} \mu(wa) \ .$$

This is the Kolmogorov extension theorem that guarantees the finite-dimensional word distributions consistently define a stochastic process [41]. We only consider shift-invariant measures and, so without loss of generality, we simply use $C_0(\omega)$ for the cylinder sets.

Crucially, the semigroup algebra and canonical machine presentation machinery for topological patterns have natural generalizations to stochastic patterns, as we now describe.

Following Ref. [42] in the context of shift spaces, a (free) *stochastic semigroup* is a function F defined on the free semigroup S, with identity element η (the empty symbol), that satisfies the following properties ([40, Definition 4.29]):

1. $F(\eta) = 1$,
2. $F(s) \geq 0$ for each $s \in S$ and $F(a) \geq 0$ for each $a \in \mathcal{A}$, and
3. $\sum_{a_i \in \mathcal{A}} F(a_i s) = \sum_{a_i \in \mathcal{A}} F(s a_i)$ for each $s \in S$.

For a sofic shift $\mathcal{X}$, we define a stochastic semigroup on $\mathcal{X}$ using the semigroup G defined above, with absorbing element e corresponding to forbidden words. We simply set $F(e) = 0$. Then, a shift-invariant measure μ satisfying Kolmogorov extension on a sofic shift $\mathcal{X}$ with semigroup G forms a stochastic semigroup as follows. For all elements of the free semigroup $a_0 a_1 \cdots a_k$—i.e., for every word $\omega = a_0 a_1 \cdots a_k$—define F such that $F(\eta) = 1$, $F(\omega) = 0$ if and only if $\omega \in \mathcal{F}$ (equivalently, $a_0 a_1 \cdots a_k = e$ in the semigroup G), and otherwise $F(\omega) = \mu\big(C_0(\omega)\big)$. Such a measure μ is called a *sofic measure* [40].

In this way, sofic measures allow for statistical structure on top of a sofic shift $\mathcal{X}$, while maintaining an algebraic structure related to $\mathcal{X}$'s semigroup algebra. More importantly, there is a canonical machine presentation associated with sofic measures, analogous to the canonical machine presentation of sofic shifts. As in the topological case for sofic shifts, the canonical machine presentation of sofic measures provides the mathematical formulation of *statistical patterns*.

Recall that the future cover semiautomaton—the canonical topological machine presentation—is defined from Eq. (1)'s future equivalence relation. The canonical stochastic machine presentation is defined through a stochastic generalization of future equivalence, called *predictive* or *causal* equivalence, defined on semi-infinite words. Each index i partitions a sequence into a semi-infinite *past* $\overleftarrow{x}_i = \{x_j\}, j \leq i$, and semi-infinite *future* $\overrightarrow{x}_i = \{x_k\}, k > i$. In the topological case, two pasts are considered future-equivalent if they have the same future—the same *set* of futures that follow. In the stochastic case, two pasts are *predictively* or *causally equivalent* if they have the same *distribution* $\Pr(\overrightarrow{X}|\overleftarrow{X})$ over futures conditioned on the past:

$$\overleftarrow{x}_i \sim_\epsilon \overleftarrow{x}_j \iff \Pr(\overrightarrow{X}|\overleftarrow{X} = \overleftarrow{x}_i) = \Pr(\overrightarrow{X}|\overleftarrow{X} = \overleftarrow{x}_j). \quad (2)$$

Just as the future equivalence relation $\sim_F$ defines the unique minimal semiautomaton presenting a sofic shift, the causal equivalence relation $\sim_\epsilon$ defines the unique minimal *hidden Markov chain* (HMC) that presents a sofic measure and its stationary stochastic process [18, 43].

Speaking simply, a hidden Markov chain $(\Xi, \mathcal{A}, \mathcal{T})$ is a semiautomaton whose deterministic symbol-labeled transitions $\mathcal{M} = \{M_0, M_1, \ldots, M_{(n-1)}\}$ are replaced by symbol-labeled transition probabilities $\mathcal{T} = \{T^0, T^1, \ldots, T^{(n-1)}\}$, where $T^a_{\xi,\xi'}$ is the probability of transitioning from state ξ to ξ' on the symbol $a \in \mathcal{A}$.

Paralleling the topological setting, the canonical stochastic machine presentation is a hidden Markov chain whose internal states Ξ—the *predictive* or *causal states*—are the equivalence classes of Eq. (2)'s causal equivalence relation. The symbol-labeled transitions are then defined through the one-step conditional distributions $\Pr(X^1 = a|\overleftarrow{X} = \overleftarrow{x}_i)$. For a given causal state ξ, we write $\Pr(\overrightarrow{X}|\overleftarrow{X} = \xi)$ since by definition each past $\overleftarrow{x}_i$ in the equivalence class $\xi = [\overleftarrow{x}_i]_\epsilon$ has the same predictive distribution.

The transition probability $T^a_{\xi,\xi'}$ is then given as $\Pr(X^1 = a|\overleftarrow{X} = \xi)$, with $\xi' = [\overleftarrow{x}_i a]_\epsilon$, where $\overleftarrow{x}_i a$ is the new past given by concatenating the observed symbol a onto the current past $\overleftarrow{x}_i$. This follows from *unifilarity* [18]: in the stochastic setting each internal state $\xi \in \Xi$ and symbol $a \in \mathcal{A}$ there is at most one internal state ξ' such that $T^a_{\xi,\xi'} > 0$. It then follows that $[\overleftarrow{x}_i a]_\epsilon$ is the same for all $\overleftarrow{x}_i \in \xi$.

Historically, the canonical stochastic machine presentation $(\Xi, \mathcal{A}, \mathcal{T})$ is known as the *ϵ-machine* [18] of the associated stochastic process. As described shortly, the

ϵ-machine is a generator of its associated statistical field theory. It generates all words with their corresponding probabilities. Similar to its topological counterpart, unifilarity additionally elevates the ϵ-machine to a predictive presentation. Prediction in the stochastic setting means identifying the predictive distribution associated with a given past. And, by definition, an ϵ-machine's causal states carry unique predictive distributions.

Summing over all symbol-labeled transitions produces a Markov transition operator over the internal causal states: $T_\epsilon = \sum_{a\in\mathcal{A}} T^a$. This operator evolves probability distributions over the causal states, regardless of the symbols involved, and specifies an order-1 Markov process over the states [18]. The left eigenvector of T_ϵ with unit eigenvalue provides the *stationary distribution* over causal states: $\pi T_\epsilon = \pi$. For ergodic systems, as we consider here, π is unique.

We emphasize that the causal states identify a hidden, internal Markov process underlying the non-Markovian process, over the symbols in $\mathcal{A}$, specified by a sofic measure on a sofic shift. This inverts the abstract definition of sofic measures, given as factor maps on Markov processes [40]. There, a shift-invariant measure obeying Kolmogorov extension on a subshift of finite type produces a Markov chain of finite order and sofic shifts are abstractly defined as factor maps of subshifts of finite type. Hidden Markov chains are then given as the pushforward of the Markov measure along the factor map. Here, in contrast, we start with a statistical field theory supported on a sofic shift. The causal equivalence relation then identifies the underlying causal states and the Markov process defined over them.

With our given assumption of a stationary ergodic process over symbols we can use the stationary distribution π and the symbol-labeled transition operators to directly extract the word probabilities $\Pr(\omega) = \mu\big(C_0(\omega)\big)$ from the ϵ-machine presentation. First, single-symbol probabilities are given as:

$$\Pr(a) = \sum_{\xi\in\Xi} \pi(\xi) \sum_{\xi'\in\Xi} T^a_{\xi\xi'} \ . \tag{3}$$

Where $\sum_{\xi'} T^a_{\xi\xi'}$ gives the probability of observing the symbol $a \in \mathcal{A}$, conditioned on being in causal state ξ, and this is summed over all states in Ξ weighted by their stationary probabilities π. We write this compactly as:

$$\Pr(a) = \langle\pi|T^a|\mathbf{1}\rangle \ , \tag{4}$$

where $\langle\pi|$ indicates π as a row vector and $|\mathbf{1}\rangle$ is the column vector of all 1s.

The probability of a word $\omega = a_0 a_1 \ldots a_k$ is then:

$$\begin{aligned}\Pr(a_0 a_1 \cdots a_k) &= \mu\big(C_0(a_0 a_1 \ldots a_k)\big) \\ &= \langle\pi|T^{a_0} T^{a_1} \cdots T^{a_k}|\mathbf{1}\rangle \ . \end{aligned} \tag{5}$$

Recall that in the topological case, the semigroup algebra of the canonical machine presentation is given through composition of the mappings $\mathcal{M} = \{M_a\}$. Now, we see the same semigroup algebraic structure in the products of the symbol-labeled transition matrices $\mathcal{T} = \{T^a\}$. In fact, the topological structure, the set of mappings $\mathcal{M}$, is recovered from the statistical structure, the set of transitions $\mathcal{T}$, by setting all nonzero elements of each $T^a \in \mathcal{T}$ to unity and then applying future equivalence.

It must be emphasized that this last step, applying future equivalence, is essential. There may be distinct ϵ-machines—presentations of statistical patterns—that are supported on the same sofic shift—topological pattern. Said another way, statistical patterns signify distinct structure supported on topological patterns. They are not merely "adding probabilities onto" topological patterns. Formally, the symbol-labeled transition matrices $\mathcal{T}$ can represent a different semigroup than that represented by their topological counterparts $\mathcal{M}$ on which they are supported. Appendix B illustrates this distinction.

Let us briefly turn to mention the quantitative benefits of having these presentations. Using the word probabilities as just described, in addition to the underlying pattern of the sofic shift that supports the ensemble, a wide range of statistical properties of the sequence ensembles, such as correlation functions, power spectral densities, and informational properties [44, 45] can be directly determined via the ϵ-machine. For example, the *Shannon entropy rate* h_μ of an ensemble measures its degree of randomness and can be calculated as:

$$\begin{aligned} h_\mu &= \lim_{\ell\to\infty} \frac{\mathrm{H}[X_{0:\ell}]}{\ell} \\ &= \lim_{\ell\to\infty} \mathrm{H}[X_0|X_{-\ell:0}] \\ &= -\sum_{\xi\in\Xi} \pi(\xi) \sum_{a\in\mathcal{A}} \sum_{\xi'\in\Xi} T^{(a)}_{\xi\xi'} \log_2 T^{(a)}_{\xi\xi'} \ , \end{aligned}$$

where $X_{0:\ell}$ for $\ell > 0$ is the random variable for the subsequence of symbols $x_0 x_1 \cdots x_{\ell-1}$. Moreover, the Shannon information in the causal states measures a process' historical memory. For more nuanced information measures of patterns in stochastic processes and spin systems, see Refs. [24, 46–52].

III. PATTERNS IN SPACETIME

The preceding demonstrated that sofic shifts and their semiautomaton presentations provide a formulation of patterns in discrete one-dimensional systems. And, it showed how they generalize to stochastic ensembles with associated information-theoretic measures. With this, one can argue that the theory of one-dimensional patterns in discrete stationary processes, augmented with the cited extensions, is largely complete. Now, we turn to the question of how to capture patterns in higher dimensions. While patterns in this new setting are amenable to similar analysis, there are key differences, new phenomena, and open problems.

Consider now the time evolution of symbols $\mathcal{A}$ on the sites of a spatial lattice $\mathcal{L}$. A *spacetime field* $\mathbf{x} \in \mathcal{A}^{\mathcal{L}\otimes\mathbb{Z}}$ is a time series $x_0, x_1, \ldots$ of spatial configurations $x_t \in \mathcal{A}^{\mathcal{L}}$. With a d-dimensional lattice $\mathcal{L} = \mathbb{Z}^d$ a spacetime field $\mathbf{x}$ is an element of a *spacetime shift space* $\mathcal{A}^{\mathbb{Z}^{d+1}}$ [15, 17]. In what follows, upper indices on field values are spatial coordinates (e.g., at time t site i has value x^i) and lower indices are time coordinates (e.g., x_t).

Multidimensional shift spaces are notoriously difficult to study, with many simple properties being uncomputable [17]. Similarly, while sofic shifts and their canonical machine presentations provide a mathematical formalism for patterns as generalized symmetries in one dimension, generalizing to higher dimensions is not straightforward. Significantly, there is not a unique generalization for finite-state machines and regular languages in higher dimensions [16]. If there were a unique generalization, we could use the semiautomata special case as the mathematical representation of high-dimensional patterns.

As we argued, models that create a compressed representation of a system's behavior must do so by harnessing patterns—predictive regularities—in the system. Fully-discrete one-dimensional systems are ideal as there is a unique minimal presentation—the ϵ-machine—of the system and, in this case, *that predictive presentation is the pattern*. The situation is more complicated in higher dimensions. A conflict arises between useful generalized spacetime symmetries and predictive presentations that faithfully generate fields in their spacetime shift spaces.

We now outline the local spacetime generalization of predictive equivalence for constructing presentations of spacetime patterns. As with finite-state machines and regular languages, degeneracy is broken when moving to higher dimensions as the spacetime generalization is not unique. In particular, we demonstrate that the shape of local "futures" determines the algebraic properties of the resulting local presentation.

A. Local Spacetime Presentations

Since an evolving spacetime field is a time series of spatial lattice configurations, it can be interpreted as a one-dimensional shift space over an exponentially-large alphabet (of lattice configurations). While formally well defined, however, this perspective is an unwieldy basis for a mathematical formulation of patterns in spacetime. First, for space and time translation symmetries, we typically consider the idealized case of infinitely-large spatial lattices. Thus, one must work with shift spaces over an infinitely-large alphabet. Second, capturing patterns *within* the spatial configurations themselves means not treating an entire lattice configuration simply as a single symbol, as done with the one-dimensional framing. There is internal organization within the spatial configurations that is dynamically and structurally important.

Therefore, we take a local approach to generalizing machine presentations of spacetime shift spaces [13, 53]. This is implemented by identifying equivalence classes over *local* pasts that have the same future set of *local* futures in spacetime. The latter can be either topological (sets) or statistical (predictive distributions).

Motivated by the causal restrictions of local interactions, *past lightcones* in spacetime are the natural choice for local pasts when building local presentations. Formally, the past lightcone $\mathtt{L}^-$ of a site (r,t) in spacetime is the set of all field values $\mathbf{x}_{t'}^{r'}$ at previous times $(t' \leq t)$ that could influence $\mathbf{x}_t^r$ through local interactions:

$$\mathtt{L}^-(r,t) \equiv \left\{ \mathbf{x}_{t'}^{r'} : \; t' \leq t \text{ and } |r' - r| \leq c(t - t') \right\} , \quad (6)$$

where c is the finite speed of information propagation in the system. We include the *present* field value $\mathbf{x}_t^r$ in (r,t)'s past lightcone, but not in its local future.

Due to the richness and complication of multidimensional shift spaces, the following explores only topological spacetime patterns. Paralleling the one-dimensional development, local presentations of topological spacetime patterns are defined through the local analog of the future equivalence relations:

$$\mathtt{L}_i^- \sim_F \mathtt{L}_j^- \iff F(\mathtt{L}_i^-) = F(\mathtt{L}_j^-) , \quad (7)$$

where $F(\mathtt{L}_i^-) = \{\text{Local futures co-occurring with } \mathtt{L}_i^-\}$. We define co-occurring local futures more precisely shortly, as there are alternatives.

Relation (7)'s equivalence classes determine the internal states for a local spacetime presentation. The generalization of symbol-labeled transitions between the internal states is constructed in terms of spatial $[\sigma^s(\mathbf{x})]_t^r = \mathbf{x}_t^{r+s}$ and temporal $[\sigma_\tau(\mathbf{x})]_t^r = \mathbf{x}_{t+\tau}^r$ shift operators. (Note that all space-time shift operators commute.) For a past lightcone $\mathtt{L}^-(r,t)$ at spacetime site (r,t), let $\sigma_\tau^s\big(\mathtt{L}^-(r,t)\big)$ denote the action of the shift operator on all of the field values in $\mathtt{L}^-(r,t)$. That is, an entire lightcone is shifted analogously to an individual spacetime site.

The right spatial transition *fringe* is defined as the set difference between $\mathtt{L}^-(r,t)$ and $\sigma^1\big(\mathtt{L}^-(r,t)\big)$. This is the generalization of the "symbol" emitted during a rightwards move in one spatial dimension. Similarly, the left spatial transition fringe is defined as the set difference between $\mathtt{L}^-(r,t)$ and $\sigma^{-1}\big(\mathtt{L}^-(r,t)\big)$, and the forward temporal transition fringe is the set difference between $\mathtt{L}^-(r,t)$ and $\sigma_1\big(\mathtt{L}^-(r,t)\big)$ [53, Fig. 4]. For simplicity we consider spacetime fields with one spatial dimension, but the generalization to higher dimensions is straightforward.

Time and space transition fringes form the appropriate alphabet to define local spacetime presentations $(\Xi, \mathcal{A}, \mathcal{M})$, with Ξ the set of past lightcone equivalence classes, $\mathcal{A}$ the set of space and time transition fringe, and $\mathcal{M}$ the fringe-labeled state transitions [53]. Briefly, a site (r,t) in spacetime has an associated internal state $\xi(r,t)$ that is the equivalence class of the past lightcone $\mathtt{L}^-(r,t)$. Consider the neighboring site $(r+1,t)$, which similarly has an associated internal state $\xi(r+1,t) = [\mathtt{L}^-(r+1,t)]_F$. The right transition fringe

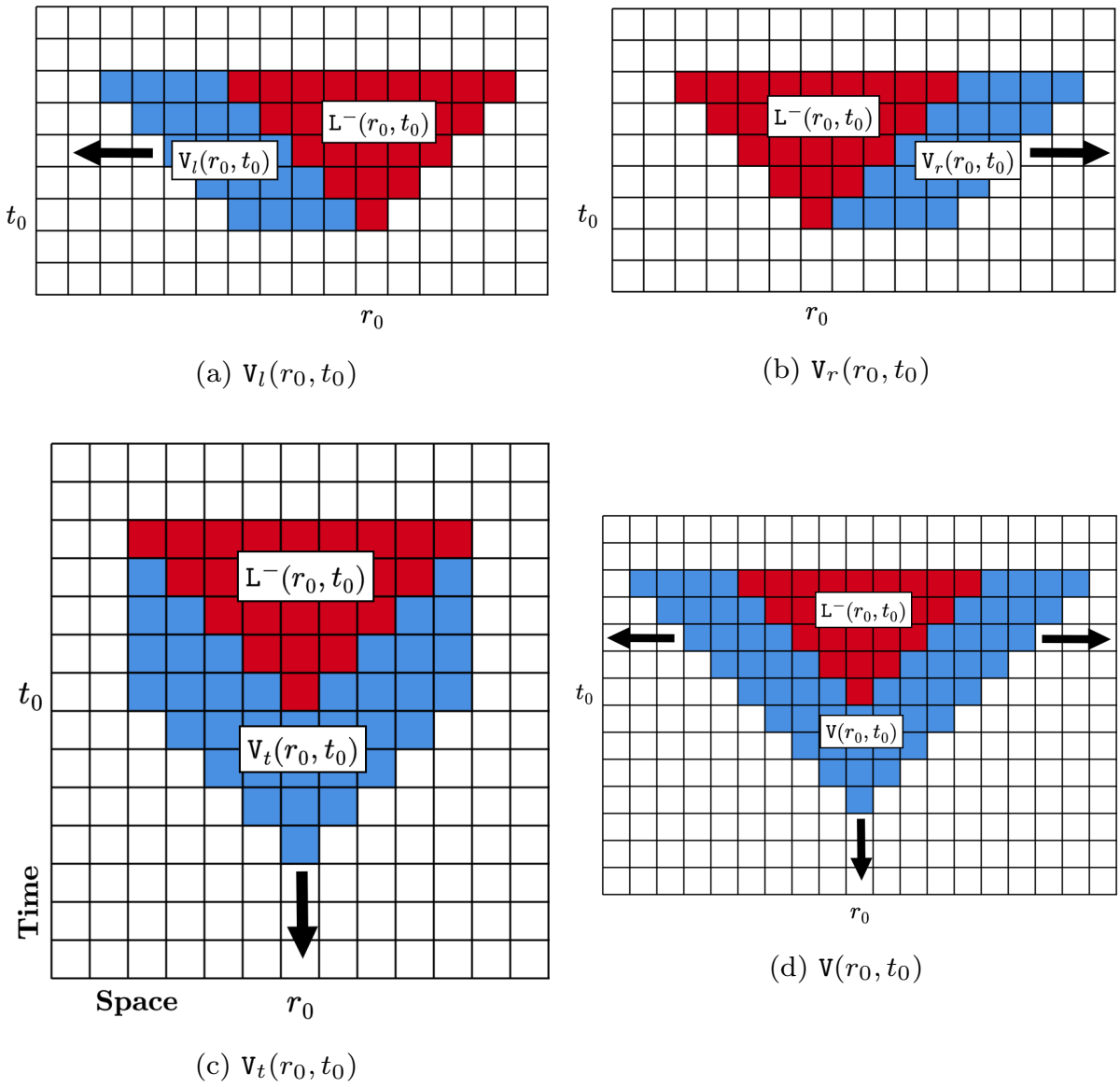


FIG. 2. Fringes induced by spacetime shifts: Co-occurring depth-4 past lightcone $\mathtt{L}^-(r_0, t_0)$ (red) and depth-4 spacetime patches resulting from concatenations of (a) left transition fringes $V_\ell(r_0, t_0)$ (blue), (b) right transition fringes $V_r(r_0, t_0)$ (blue), (c) forward transition fringes $V_t(r_0, t_0)$ (blue), and (d) unions of left, right, and forward transition fringes $V(r_0, t_0)$ (blue). Arrows indicate the direction(s) in which local spacetime patches may be generated with successive concatenations to the seed past lightcone $\mathtt{L}^-(r_0, t_0)$.

provides the missing information to construct $\mathtt{L}^-(r+1, t)$ from $\mathtt{L}^-(r, t)$. Therefore, $\xi(r, t)$ plus a right transition fringe uniquely determines $\xi(r+1, t)$. That is, such local presentations are *unifilar*, and therefore have the requisite structure of predictive presentations. Recall that predictive models are also generative.

Generating words is rather straightforward for machine presentations in one dimension, markedly less so for local spacetime presentations. The fringe-labeled transitions establish local spacetime presentations $(\Xi, \mathcal{A}, \mathcal{M})$ as local generative spacetime models. Local spacetime patches, local "words", are generated via concatenation of transition fringe "symbols". Denote concatenations of right spatial fringes generally as $\mathtt{V}_r$, to signify the particular shape of the resulting spacetime patches. Similarly, let $\mathtt{V}_l$ be the spacetime patches resulting from concatenations of left spatial fringes and $\mathtt{V}_t$ the patches from forward temporal fringes.

Note that spacetime patches are more than merely collections of values from spacetime sites. The configuration and space-time relation among the values matter. This is why we refer to their *shape*. The $\mathtt{V}_r$ patches are a distinct form of local spacetime words apart from $\mathtt{V}_l$ and $\mathtt{V}_t$. Only patches of the same shape may be concatenated together

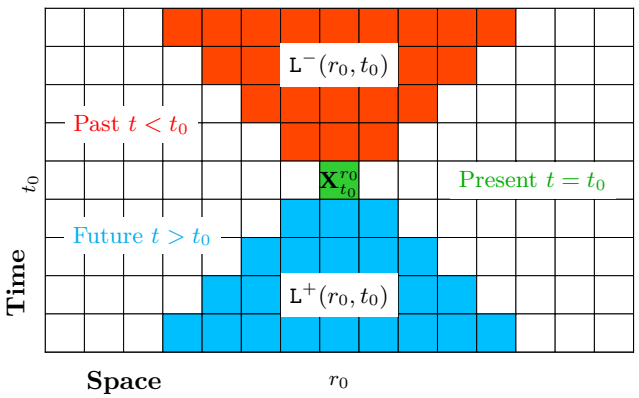


FIG. 3. Co-occurring past ($\mathtt{L}^-$) and future ($\mathtt{L}^+$) lightcones at a spacetime site (r_0, t_0) in 1+1 dimensions with $c = 1$.

to extend patches.

Let $\mathtt{V}$ be the spacetime patches resulting from the union of forward, left, and right transition fringes. Let *depth* denote the number of single-fringe symbols concatenated together in a particular spacetime patch word. See Fig. 2.

Concatenating fringes into local spacetime patches has the same semigroup algebra structure as in the one-dimensional setting. Similarly, the semigroup algebra is captured by the local presentations through the action of the fringe-labeled transitions $\mathcal{M}$, analogous to the one-dimensional presentations $\mathcal{P}(\mathcal{X})$. We now examine possible choices of local futures and the algebraic properties of the induced local presentations.

B. The Shape of Local Futures

Recall that Relation (7) did not specify local futures when defining local presentations. The future sets $F(\mathtt{L}^-)$ of past lightcones there were intentionally left ambiguous to allow for possible alternatives—alternatives that arise to address several subtleties of spacetime shifts.

Prior work assumed that the natural choice for a local future is a *future lightcone* $\mathtt{L}^+$ [13, 53, 54]. The latter is defined as all field values at subsequent times that could possibly be influenced from the given spacetime site $\mathbf{x}^r_t$ through the local interactions:

$$\mathtt{L}^+(r, t) \equiv \left\{ \mathbf{x}^{r'}_{t'} \,:\, t' > t \text{ and } |r' - r| \leq c(t' - t) \right\} . \quad (8)$$

Co-occurring past ($\mathtt{L}^-$) and future ($\mathtt{L}^+$) lightcones at spacetime point (r_0, t_0) are depicted in Fig. 3 for a $1+1$ dimensional spacetime field with $c = 1$. Local presentations constructed as equivalence classes of past lightcones that have the same future set of future lightcones are known as *local causal states*.

Interestingly, the benefit of local causal-state models derives from generalized symmetries [13, 14, 54]. However, this is a spacetime symmetry distinct from the semigroup algebra of fringe concatenations. While we describe these spacetime symmetries in more detail below, we first demonstrate that such symmetries are at odds with the semigroup algebra of fringe concatenations for local causal states. That is, while employing lightcones as local futures allows one to discover useful spacetime symmetries, the resulting local machine presentation $(\Xi, \mathcal{A}, \mathcal{M})$ is *not* a faithful generator of the underlying spacetime shift space. Even though all local presentations possess the unifilarity property of predictive models, the more basic generative ability falls short unless the appropriate local futures are used. We again emphasize that phenomena in higher dimensions are complicated—sometimes in counterintuitive ways, as the following demonstrates.

Elementary cellular automata (ECA)—see App. C—produce spacetime shift spaces with nontrivial patterns and structure in their spacetime fields; e.g., domains, particles, and particle interactions [55–57]. In addition, their one-dimensional fully-discrete spatial lattices are shift spaces and so provide a link between one-dimensional shifts and $1 + 1$ dimensional shifts of their spacetime fields.

Following convention, denote the mapping from a past lightcone to its equivalence class or, equivalently, to its associated local (causal) state $\xi_{\mathtt{L}^-}$, as $\epsilon(\mathtt{L}^-) = [\mathtt{L}^-]_F = \xi_{\mathtt{L}^-}$. Crucially, this provides a local pointwise mapping over any spacetime field $\mathbf{x}$. Each site (r, t) has a unique past lightcone $\mathtt{L}^-$ that is then mapped to its local state via $\epsilon(\mathtt{L}^-)$. We use this pointwise mapping to transform a spacetime field $\mathbf{x}$ to an associated local-state field $\mathcal{S} = \epsilon(\mathbf{x})$ that shares the same coordinate geometry as $\mathbf{x}$. Each site $\mathcal{S}^t_r$ in the local state field $\mathcal{S}$ is the local state $\mathcal{S}^t_r = \xi^-_{\mathtt{L}} = \epsilon(\mathtt{L}^-)$ that is the image under the ϵ-map of the past lightcone $\mathtt{L}^-$ at the spacetime point $\mathbf{x}^r_t$.

An example spacetime field $\mathbf{x}$ from ECA Rule 90 is shown in Fig. 4 consisting of black ($x^r_t = 1$) and white ($x^r_t = 0$) squares. The associated local causal-state field $\mathcal{S}$ is shown as the overlaid colored letters, using lightcones as local futures. In the case of Rule 90, there is a single local causal state, labeled as A—all past lightcones are future-lightcone equivalent.

This is consistent with prior findings [13, 54] that connect the local causal states with the canonical machine presentations of one-dimensional sofic shifts that represent *invariant sets* of spatial configurations for the ECA. The full 2-shift is invariant under Rule 90, meaning there are no forbidden words in the one-dimensional spatial configurations generated by Rule 90. All binary strings have pre-images under the global dynamic Φ of Rule 90. Recall that the canonical machine presentation of full shifts consists of a single internal state and, hence, we expect the single local causal-state for the spacetime fields produced by Rule 90.

Even though there are no forbidden words in Rule 90's one-dimensional spatial configurations, there certainly *are* forbidden patches in its spacetime fields. In particular, note that the ECA's local update rule ϕ corresponds to a spacetime patch in the shape of a depth-1

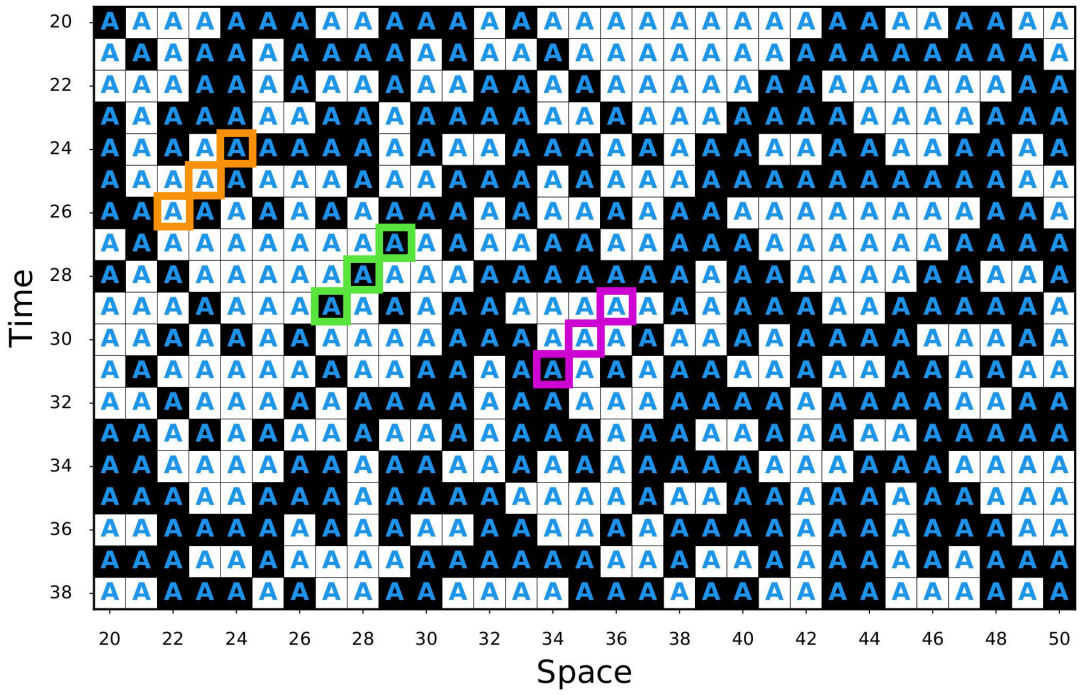


FIG. 4. ECA Rule 90 spacetime field depicted as white (0) and black (1) squares. The corresponding local causal-state field is overlaid with colored letters; simply the single causal state A. Three sample right-transition fringes for past lightcone depth-2 are highlighted in colored (orange, green, purple) boxes.

past lightcone—the present spacetime site together with its local neighborhood one time-step in the past. Therefore, any spacetime patch of the same shape that is inconsistent with local update rule ϕ is forbidden in the spacetime shift space of the ECA's spacetime fields. For example, the following patch is allowed by Rule 90, since $\phi_{90}(010) = 0$:

$$\begin{array}{ccc} 0 & 1 & 0 \\ & 0 & \end{array}$$

Therefore:

$$\begin{array}{ccc} 0 & 1 & 0 \\ & 1 & \end{array}$$

is a forbidden spacetime patch in Rule 90's spacetime shift space.

From ECA Rule 90's single local causal-state it is easy to see that the local causal-state machine presentation is not a consistent spacetime generator. Since there is only one state, A, all fringe-labeled transitions occur from state A back to itself. Thus, all fringe symbols that occur may be concatenated together to form spacetime patches. Figure 4 highlights three right-transition fringes. If they are concatenated in order of orange, green, magenta (left to right) they form the depth-3 $\mathtt{V}_r$ patch:

$$\begin{array}{ccc} & 1 & 1 & 0 \\ 0 & 1 & 0 & \\ 0 & 1 & 1 & \end{array}$$

This patch contains the spacetime word given above that is forbidden in Rule 90's spacetime shift. In fact, the other word of the same shape in this patch is also forbidden by Rule 90:

$$\begin{array}{ccc} 1 & 1 & 0 \\ & 0 & \end{array} \quad ,$$

since $\phi_{90}(110) = 1$.

Concatenations of left and forward fringes produce, respectively, $\mathtt{V}_l$ and V_t spacetime patches that similarly contain words forbidden by Rule 90's spacetime shift space, as shown numerically in a supplementary Jupyter Notebook [58]. Similar results for another ECA shift space, the domain of Rule 18 discussed in more detail below, are also given in a supplementary Jupyter Notebook [59]. Empirically, we have found that local causal state presentations are generically not faithful generative models of their CA spacetime shift spaces.

Comparing Fig. 3 with the $\mathtt{V}_i$ patch shapes in Fig. 2, there is little to no overlap between the "futures" (lightcones in this case) used to define the equivalence classes of past lightcones and the spacetime patches generated from fringe-labeled transitions between the equivalence classes. Therefore, it is not surprising that local presentations created with future lightcones are not faithful generators of their spacetime shifts.

In one dimension, there is only one "future" that always follows from the past. When translating forward, what once was a future becomes part of the past. This

is not necessarily the case for local pasts and futures in spacetime. The predictive ability of unifilar models follows from this succession of futures becoming pasts, when combined with future equivalence. As we now see, future equivalence without the successional relation between local pasts and futures yields models that, while unifilar, are not consistent spacetime generators.

Constructively, this insight points the way to creating local presentations that *are* faithful generators. As the future lightcones do not play a role in generating local spacetime patches, they should not be used to define past lightcone equivalence in Relation (7). Rather, we should use spacetime patch shapes—denote them $\mathtt{V}_i$—that are to be generated as our notion of a local future for defining future equivalence in Relation (7). Local causal states are defined using the future sets with future lightcones $F(\mathtt{L}_i^-) = \{\mathtt{L}^+ \text{ co-occurring with } \mathtt{L}_i^-\}$. We can alternatively define local machine presentations whose internal states are defined as equivalence classes from Relation (7) using $F(\mathtt{L}_i^-) = \{\mathtt{V}_i \text{ co-occurring with } \mathtt{L}_i^-\}$.

Indeed, as demonstrated computationally in the supplementary Jupyter Notebook Ref. [58], using $\mathtt{V}_r$ shapes with $F(\mathtt{L}_i^-) = \{\mathtt{V}_r \text{ co-occurring with } \mathtt{L}_i^-\}$ yields a local machine presentation that is a faithful generator of $\mathtt{V}_r$ spacetime patches. However, this presentation is *not* a faithful generator of $\mathtt{V}_l$ or $\mathtt{V}_t$ spacetime patches. Similar to the use of future lightcones, local presentations can only generate faithful spacetime patches in the shape used to define $F(\mathtt{L}^-)$. Therefore, as also demonstrated in Ref. [58], using $\mathtt{V}_l$ to define $F(\mathtt{L}^-)$ results in local presentations that faithfully generate $\mathtt{V}_l$, but not $\mathtt{V}_r$ nor $\mathtt{V}_t$. Similar for $\mathtt{V}_t$.

This motivates the definition of $\mathtt{V}$ in Fig. 2 (d) as the union of $\mathtt{V}_r$, $\mathtt{V}_l$, and $\mathtt{V}_t$. Local presentations defined using $F(\mathtt{L}_i^-) = \{\mathtt{V} \text{ co-occurring with } \mathtt{L}_i^-\}$ are faithful generators of spacetime patches in *all* directions emanating from a seed past lightcone. As the resulting presentations are also unifilar, additionally they are consistent predictive models of local spacetime patches. These $\mathtt{V}$ presentations can thus be seen as the most natural generalization of the predictive canonical machine presentations $\mathcal{P}(\mathcal{X})$ for one-dimensional sofic shifts. Moreover, they possess the same semigroup algebra of "symbol" (fringe) concatenation that allows the generation and prediction of arbitrarily-long "words"—arbitrarily large spacetime patches.

That said, as the following now details, local causal-state presentations that use future lightcones to define $F(\mathtt{L}^-)$ possess interesting and useful generalized spacetime symmetries that are lost when using $\mathtt{V}$ presentations.

C. Generalized Spacetime Symmetries

Recall from above that machine transitions in one-dimension determine the structural relations among internal states and that they are driven by the shift operator. We just saw that the analogous fringe-labeled transition structure in spacetime is similarly related to spacetime shift operators through the definition of the fringes. However, the spacetime shift operators provide structural algebraic relations among local states *independent of* fringe symbols and their concatenated spacetime words. This is due to the local pointwise ϵ-map and the resulting shared coordinate geometry between spacetime fields $\mathbf{x}$ and corresponding local state fields $\mathcal{S} = \epsilon(\mathbf{x})$.

To see this, consider two space-adjacent sites in a spacetime field, $\mathbf{x}_t^r$ and $\mathbf{x}_t^{r+1} = \sigma^1 \mathbf{x}_t^r$. Under the ϵ-map, this produces $\xi_i = \mathcal{S}_t^r$ and $\xi_j = \mathcal{S}_t^{r+1} = \sigma^1 \xi_i$. As past lightcones are defined solely in terms of spacetime distances, they are equivariant under spacetime isometries. In the present setting these are translations.

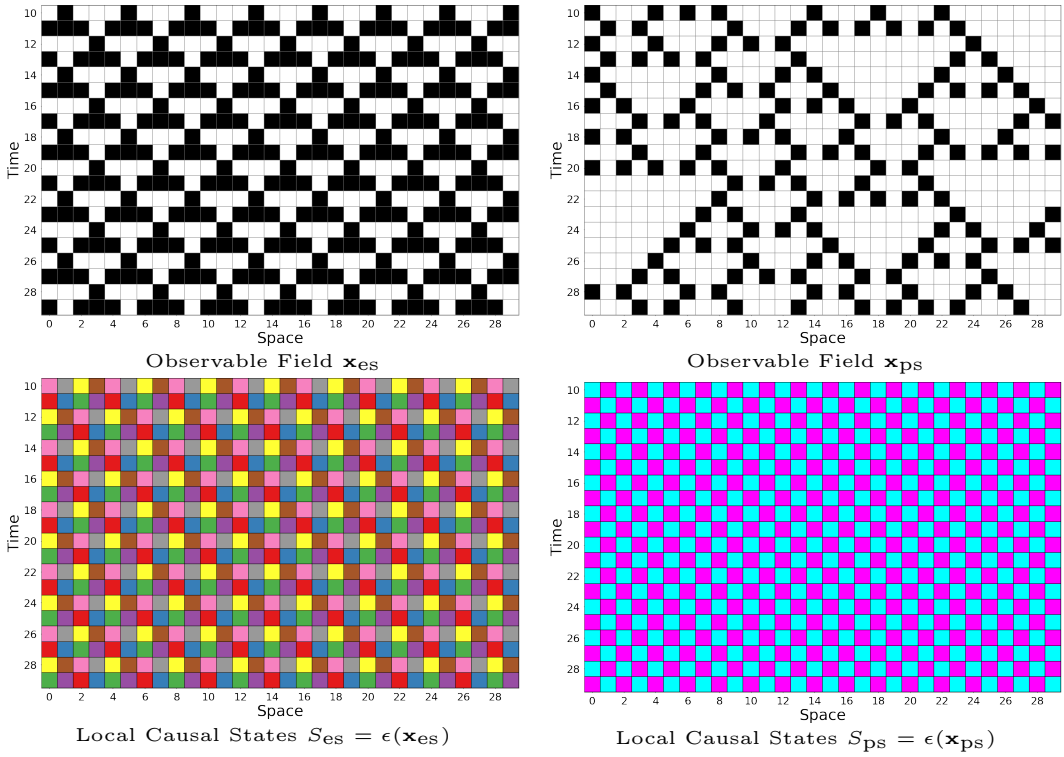
For example, if a spacetime field $\mathbf{x}$ has *exact* time and space translation symmetries $\sigma_\tau \mathbf{x} = \mathbf{x}$ and $\sigma^s \mathbf{x} = \mathbf{x}$, for some s and τ, then the corresponding local state field $\mathcal{S} = \epsilon(\mathbf{x})$ shares these symmetries: $\sigma_\tau \mathcal{S} = \mathcal{S}$ and $\sigma^s \mathcal{S} = \mathcal{S}$. This is because the spacetime shift operators act equivalently on lightcones, so that $\sigma^s \mathtt{L}_i^- = \mathtt{L}_i^-$ and $\sigma_\tau \mathtt{L}_i^- = \mathtt{L}_i^-$. Therefore, $\epsilon\big(\sigma^s \mathtt{L}_i^-\big) = \epsilon(\mathtt{L}_i^-)$ and $\epsilon\big(\sigma_\tau \mathtt{L}_i^-\big) = \epsilon(\mathtt{L}_i^-)$. Note that this argument is independent of which local future shape is used, as they all define a local ϵ-map on past lightcones.

When future lightcones specifically are used as local futures, we previously showed (i) there are spacetime fields $\mathbf{x}$ that do not have translation symmetries, but (ii) the associated local causal state field $\mathcal{S} = \epsilon(\mathbf{x})$ *does* [13, 54]. These are the spacetime generalizations of partial and hidden symmetries; cf. Fig. 1.

For the cellular automata examples in Refs. [13, 54] and shown here in Fig. 5, generalized symmetries—exact, partial, and hidden—in the spacetime fields are generated by the evolution of invariant one-dimensional sofic shifts (spatial configurations) under the CA dynamic. Such spacetime regions are known as *domains* [55–57]. Interestingly, exact symmetry domains are generated from the evolution of exact symmetry sofic shifts. Moreover, stochastic partial symmetry domains are generated from stochastic partial symmetry sofic shifts and hidden symmetry domains from hidden symmetry shifts. Examples of each of these cases are shown in Fig. 5 as the generalizations of the one-dimensional cases in Fig 1. Appendix D displays the presentations $\mathcal{P}(\mathcal{X})$ for each sofic shift $\mathcal{X}$ used to generate the fields in Fig. 5.

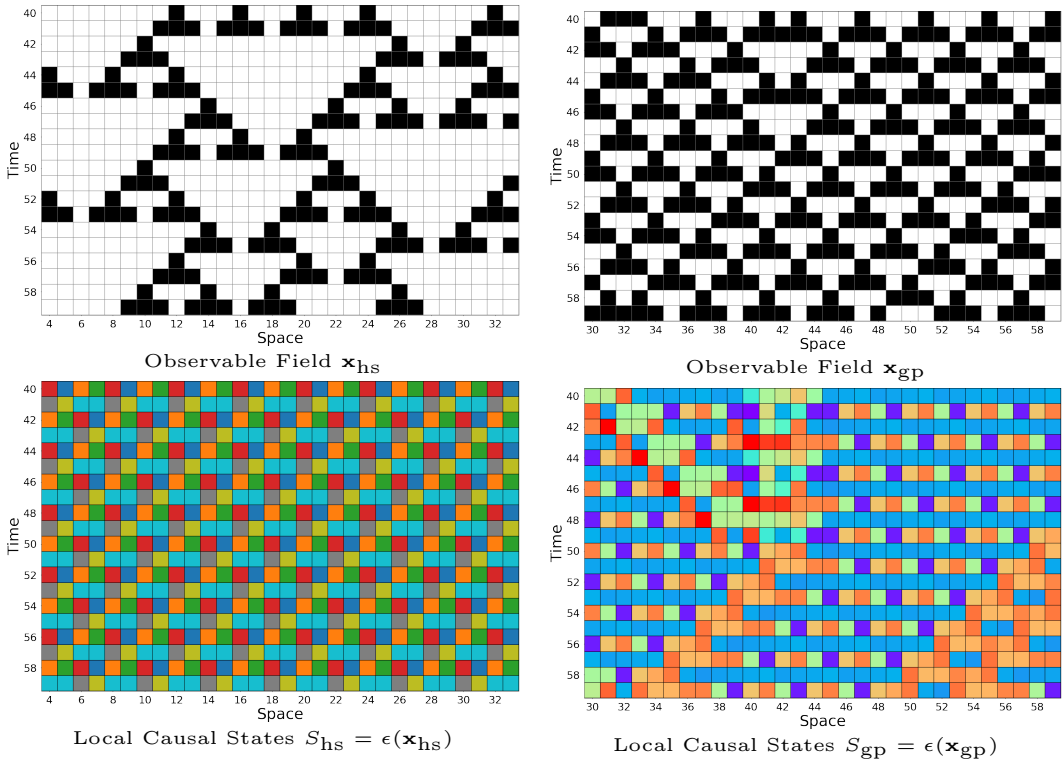
Exact symmetries are straightforward, as just described. There is a finite s and τ such that for *every* site (r, t) in spacetime we have that $\sigma_\tau \mathbf{x}_t^r = \mathbf{x}_t^r$ and $\sigma^s \mathbf{x}_t^r = \mathbf{x}_t^r$. Due to the shared coordinate geometry and isometry equivariance of past lightcones, this applies for the local causal state field as well: $\sigma_\tau \mathcal{S}_t^r = \mathcal{S}_t^r$ and $\sigma^s \mathcal{S}_t^r = \mathcal{S}_t^r$. The exact symmetry field shown in Fig. 5 (a) is a sample of the domain of ECA Rule 54. As can be seen, both the spacetime fields and corresponding local state fields have a period-4 translation symmetry in both time and space.

For partial symmetries, some, but not all, of the spacetime coordinates return to themselves after fixed trans-



(a) Exact Symmetry

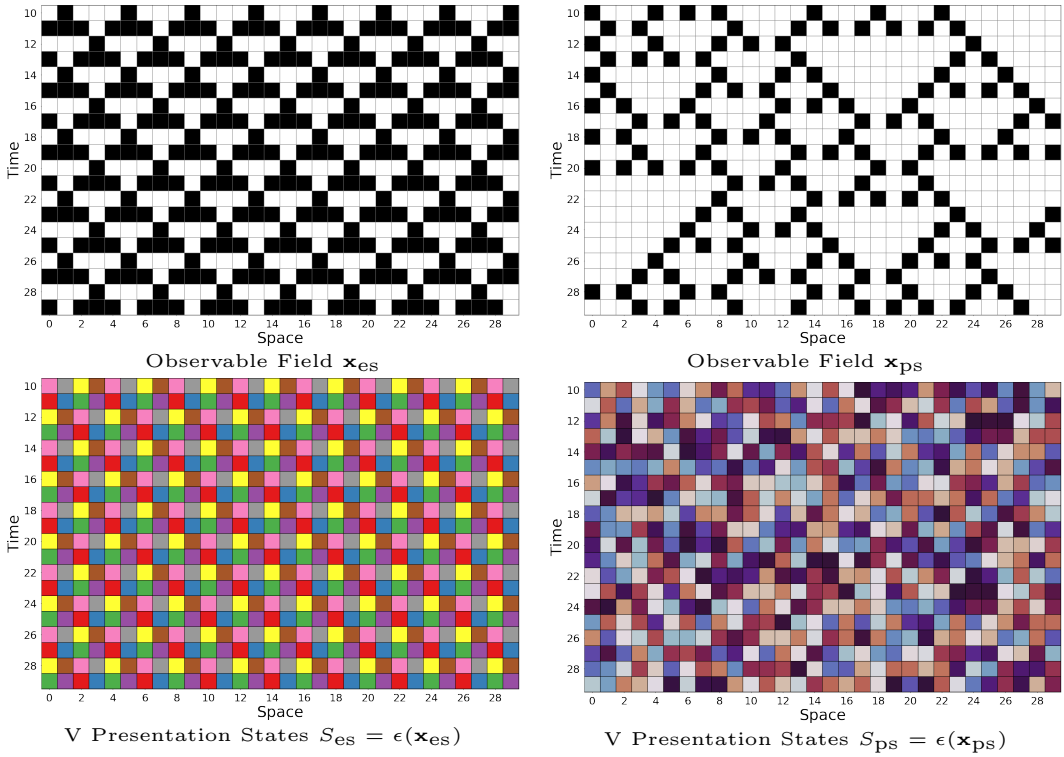
(b) Partial Symmetry



(c) Hidden Symmetry

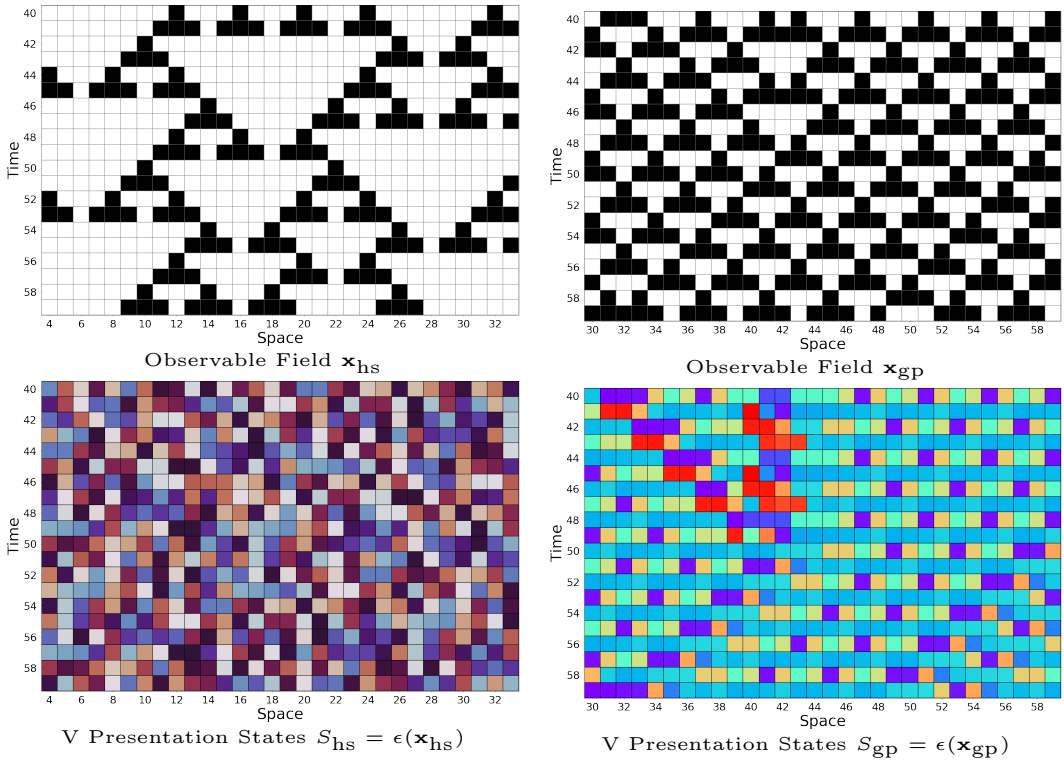
(d) General Pattern

FIG. 5. Spacetime pattern classes: Spacetime fields $\mathbf{x}$ (above) and corresponding local causal state fields $\mathcal{S} = \epsilon(\mathbf{x})$ (below) for (a) an exact symmetry (ECA Rule 54 domain), (b) a partial symmetry (ECA Rule 18 domain), (c) a hidden symmetry (ECA Rule 22 domain), and (d) a general pattern (ECA Rule 54 evolving random initial configuration).



(a) Exact Symmetry

(b) Partial Symmetry



(c) Hidden Symmetry

(d) General Pattern

FIG. 6. Spacetime fields $\mathbf{x}$ shown in Fig. 5 (above) and corresponding V state fields $\mathcal{S} = \epsilon(\mathbf{x})$ (below) for (a) an exact symmetry, (b) a partial symmetry, (c) a hidden symmetry, and (d) a general pattern.

lation, just as in one dimension. For example, there are some sites (r^*, t^*) such that $\mathbf{x}_{t*}^{r*+s} = \sigma^s \mathbf{x}_{t*}^{r*} = \mathbf{x}_{t*}^{r*}$, but this does not hold for all (r, t). As in one-dimension, a special case of partial symmetries are stochastic symmetries that become symmetric after a wildcard substitution. For example, consider a 1+1 dimensional spacetime field with $\mathcal{A} = \{0, 1\}$ such that it has a checkerboard layout with 0s on the black squares and wildcards (either 0 or 1) on the white squares. The partial (stochastic) symmetry example shown in Fig. 5 (b) is ECA Rule 18's domain, which is a strict subset (subshift) of the 0-wildcard checkerboard shift space. We examine these two spacetime shift spaces in more detail below.

Unlike in one-dimension, hidden symmetries in higher dimensions correspond to exact symmetries in the local causal-state field for spacetime fields that themselves have no symmetries, exact or partial. The hidden symmetry spacetime fields in Figs. 5 (c) and 6 (c) are samples of ECA Rule 22's domain and their structure is manifestly harder to detect from visual inspection.

What is most important to note here is that the observable field $\mathbf{x}$ does not have any space or time translation symmetries, exact or partial. The corresponding local causal-state field $\mathcal{S} = \epsilon(\mathbf{x})$ (shown in Fig. 5), though, does have symmetries that are period-4 in both time and space. Essentially, there are motifs that appear in $\mathbf{x}$—such as, the black "triangles" that occur with a local period-4 structure (e.g., 1110 in space and 1100 in time). However, in contrast to a stochastic symmetry field, there is no global symmetry that captures how these stochastic motifs occur. The global hidden symmetry is more complicated than can be revealed by simple wildcard substitutions. It is uncovered, though, by the local causal-state field. This structure, and its relation to the local causal states, is thoroughly detailed in Ref. [54].

A general spacetime pattern then is one for which the local causal-state field does not exhibit spacetime symmetries. There is still an algebraic relation among the local states from the spacetime shift operators, but they do not correspond to spacetime symmetries.

Note that the example in Fig. 5 (d) has regions that are *locally* symmetric in both $\mathbf{x}$ and $\mathcal{S}$. However, this symmetry is globally broken by localized defects or coherent structures [13]. Figure 5 (d) is produced from from ECA Rule 54 evolving a random initial configuration. The local symmetry regions are instances Rule 54's domain—the exact symmetry field shown in Fig. 5 (a).

We emphasize that, with the exception of exact symmetries, these generalized spacetime symmetries are empirically observed *only* with local causal-state presentations that employ future lightcone shapes. For example, the corresponding local state fields of presentations using `V` local futures is shown in Fig. 6 with the same spacetime fields as Fig. 5. As expected, the exact symmetry case also has space-time translation symmetries in the local state field in (a). However, there are no space-time symmetries present in the local state fields for partial (b) or hidden (c) symmetry spacetime fields when `V` local futures are employed.

Altogether these observations portray a somewhat unsatisfactory scenario. On the one hand, lightcone local futures and the resulting local causal-state models produce insightful and predictive generalized symmetries in spacetime. Moreover, these clearly connect to the one-dimensional sofic shifts of CA spatial configurations, as demonstrated in Refs. [13, 54]. However, we just demonstrated that local causal-state models are unfaithful generators of their underlying CA shift spaces.

On the other hand, we also showed how to create faithful local generators using `V` future shapes. These presentations thus appear to be the most natural local generalization of the canonical machine presentations in one dimension. In this case, though, the useful spacetime generalized symmetries are lost with these generative models.

Let us now examine in more detail the trade-offs between these two local presentations by diving deeper into the case of stochastic symmetries.

IV. CASE STUDY: STOCHASTIC SYMMETRIES

The partial (stochastic) symmetry spacetime fields in Figs. 5 (b) and 6 (b) are generated from ECA Rule 18 evolving a string from its domain's invariant sofic shift. Appendix D shows that these are points in the stochastic symmetry sofic shift with the form 0-Σ, where Σ is a wildcard that can be either 0 or 1.

It is easy to see from Rule 18's lookup table ϕ that the wildcard locations oscillate each time step between even- and odd-indexed lattice sites. Thus, the spacetime field can be interpreted as a checkerboard pattern of fixed 0s and wildcards $\Sigma = \{0, 1\}$; giving the checkerboard pattern seen in the local causal-state field shown in Fig. 5 (b). In this way, the two local causal states can be interpreted as a "fixed" 0 state and a wildcard state. The local states of the `V` machine presentation, though, do not occur in a checkerboard pattern in Fig. 6 (b). And so, they clearly cannot be assigned such semantic labels.

While the fixed-0 and wildcard semantic labels are appealing for the two local causal states, they are also misleading. For a given spatial configuration, the fixed-0 and wildcard semantics are appropriate, as it describes the (invariant) one-dimensional sofic shift of the spatial configurations. However, it is again easy to see from the Rule 18 lookup table that the wildcard semantics can no longer be assigned to the full spacetime field. Specifically, in spacetime the local update rule ϕ_{18} forbids certain spacetime patches of the form:

$$\begin{array}{ccc} & \Sigma & \\ \Sigma & 0 & \Sigma \end{array}$$

Therefore, the spacetime shift space of Rule 18's domain is a proper subset (subshift) of the 0-Σ checkerboard shift

space for which all realizations of the above spacetime patch are allowed.

Figure 7 demonstrates that both the local causal states and the V machine presentation local states reveal a checkerboard symmetry for spacetime fields in the 0-Σ shift space. In this case, *both* types of local states certainly *do* carry the semantics of fixed-0 and wildcard Σ in spacetime.

What does this say then about Rule 18's domain in the context of the alternative local-state presentations in Figs. 5 (b) and 6 (b)?

First, it is interesting and not entirely clear why the choice of lightcone local futures produces local causal states that carry the strictly-spatial semantics of fixed-0 and wildcard Σ. Especially so, when we know the wildcard semantics do not hold in spacetime. Whatever the reason, it is key to interpreting spacetime patterns through the local causal-state fields, including identifying coherent structures as locally-broken generalized symmetries in spacetime [13].

Second, this implies that the V machine presentation states are a more nuanced representation of the spacetime pattern of Rule 18's domain. From above, we know the non-fixed-0 sites cannot be strictly interpreted as wildcards in spacetime. However, we can interpret these as contextually-constrained wildcards. From the 0-Σ patch above, the outcome of the bottom Σ is constrained by ϕ_{18} and the outcomes of the preceding Σs. Propagating these constraints through space and time clearly becomes complicated very quickly. However, this is a more appropriate semantic interpretation of Rule 18's domain. And so, this seems to be to what the V machine states correspond. As shown in a supplementary Jupyter Notebook [60], the local state field displayed in Fig. 6 (b) is reconstructed from past lightcones of depth-8 and Vs of depth-3. This produces 767 local V states to (approximately) capture the spacetime pattern of Rule 18's domain. This is in contrast to the three local causal states reconstructed with the same past and future depths. (For finite-depth past lightcones, there is a third "indeterminate" state for the all-0 past lightcone.)

Note that the fixed-0 semantics still holds in spacetime and that this is not captured by the V presentation states. Given that V machine presentations are constructed to be faithful local generators of their spacetime shift spaces, it seems there is a similar spacetime contextuality to the fixed-0 sites that is necessary for faithful local generation. The spacetime contextuality present in the V presentation states is absent in the local causal states. And, it is the latter that reveals the spacetime symmetries observed in Fig. 5.

Taken all together, the spacetime patterns of partial and hidden symmetry ECA spacetime shift spaces are exceedingly complicated. Local causal-state presentations constructed from future lightcones and the V machine presentations capture different aspects of these patterns. One the one hand, V machine presentations are faithful generators and predictors of the spacetime patterns and so capture the patterns in a more direct connection with the one-dimensional case. However, local causal states capture generalized spacetime symmetries that, in the case of ECA domain shift spaces, closely connect with the (evolution of) the one-dimensional shift spaces of the invariant spatial configurations that, in turn, possess the corresponding generalized symmetries.

V. CONCLUSIONS

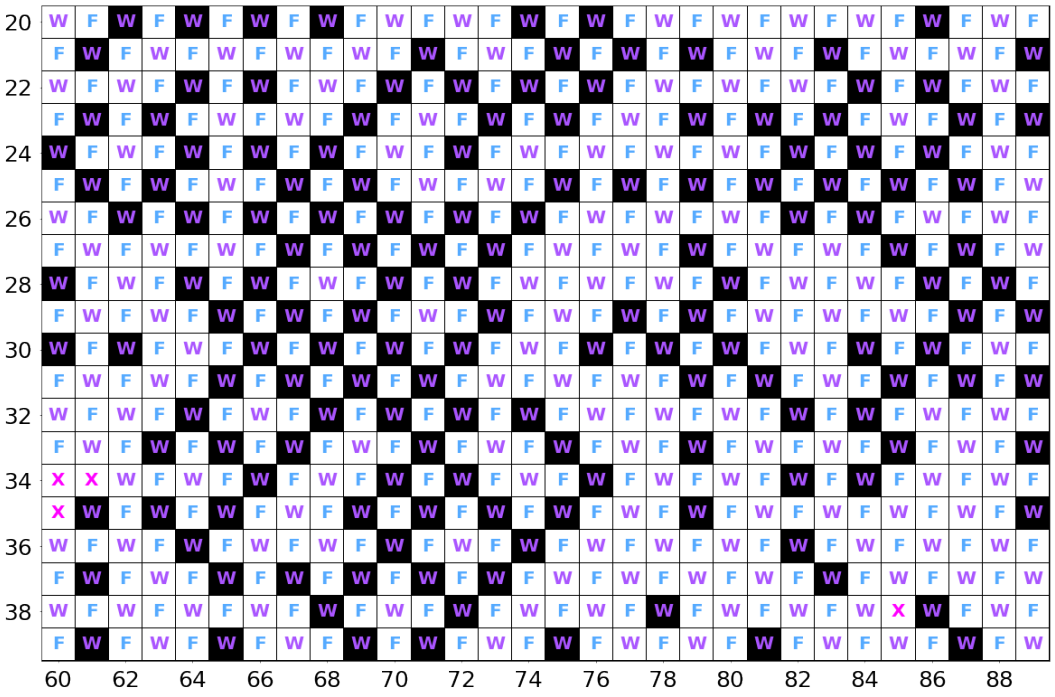
Building on Refs. [43] and [18]'s foundations, there is a growing body of results that use machine presentations defined from future or predictive equivalence to discover inherent, often hidden, pattern and structure. This includes pattern and structure in disordered crystals [61], thermodynamic environments [62], atmospheric turbulence [63], bacteria behavior [64], controlled quantum processes [65], and more.

The first half of the preceding development synthesized the arguments for sofic shifts and their machine presentations as mathematical formulations of pattern and structure. In particular, the manner in which they generalize the perfect regularity of exact symmetries and the associated group algebra has been rigorously clarified. It also connected to recent results on sofic measures and their relation to stochastic ϵ-machine presentations of statistical patterns supported on sofic shifts.

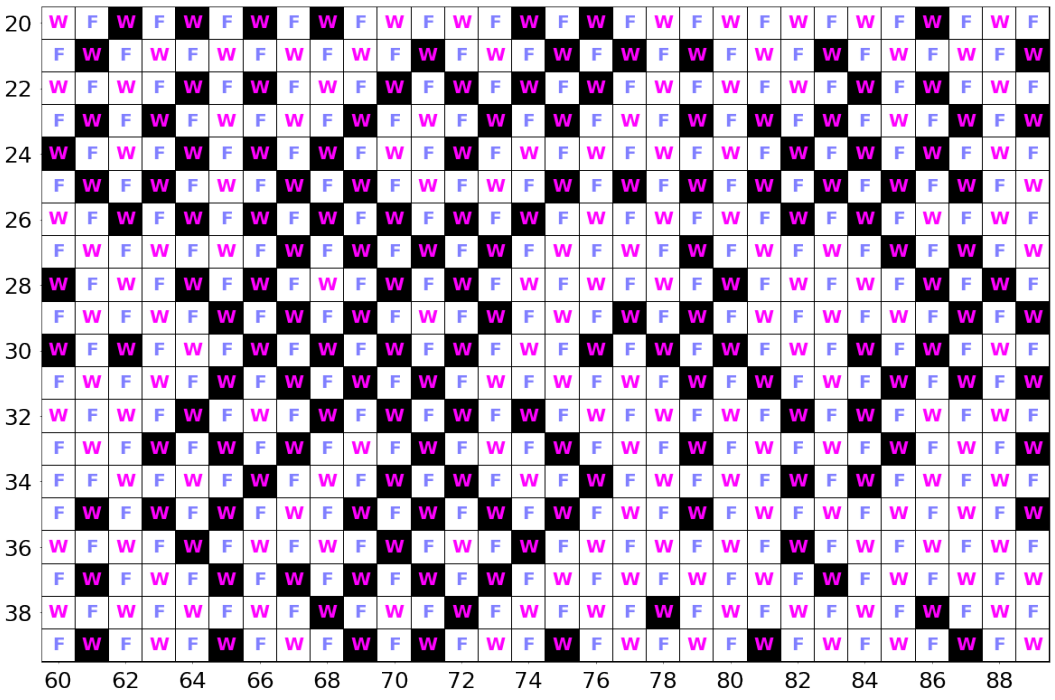
The development's second half overviewed the local approach to spacetime machine presentations. We showed that the standard local causal-state approach that uses lightcones as local futures reveals useful generalized spacetime symmetries. However, this comes at a cost: it does not lead to faithful generative models of the underlying shift space. This motivated introducing an alternative local presentation model using spacetime V shapes as local futures. We showed that these presentations are faithful generative models and that they do not possess the same generalized spacetime symmetries of the local causal states. The seeming mutual-exclusion of multiple local presentations is emblematic of the difficulty of high-dimensional shift spaces. The novel constructions given here—in particular the generative local presentations using V futures—may provide new paths of inquiry for investigating the organization of these rich and challenging spaces.

One advantage of the nongenerative approach using local predictive equivalence over future lightcones is that it does not rely on a finite alphabet for labeled transitions to provide the algebraic structure among local causal states. Thus, local causal states are well-defined and can be algorithmically approximated for continuum field theories. For example, they have been used to extract coherent structures in complex fluid flows [14]. The results presented here provide a theoretical underpinning and a "physics of organization" behind these unsupervised physics-informed machine learning algorithms.

Supporting Python code and supplementary Jupyter



(a) Local causal states



(b) V presentation local states

FIG. 7. Sample field from 0-Wildcard shift space in black (1) and white (0) squares with (a) local causal states and (b) V presentation local states (b) overlaid. In both cases, the local states can be assigned fixed-0 and wildcard semantics. And so, they are labeled F and W, respectively. The local causal states in (a) have an additional "indeterminate" state assigned to the all-0 past lightcone, labeled as X; see, for example, the field at $r = 60$ and $t = 34$).

Notebooks can be found at `https://github.com/adamrupe/ca_patterns`.

ACKNOWLEDGMENTS

The authors thanks Mikhael Semaan for helpful comments and discussions. The authors also thank the Telluride Science Research Center for hospitality during visits and the participants of the Information Engines Workshops there. AR acknowledges the support of the U.S. Department of Energy through the LANL/LDRD Program and the Center for Nonlinear Studies. JPC acknowledges the kind hospitality of the Santa Fe Institute, Institute for Advanced Study at the University of Amsterdam, and California Institute of Technology for their hospitality during visits. This material is based upon work supported by, or in part by, FQXi Grant number FQXi-RFP-IPW-1902 and U.S. Army Research Laboratory and the U.S. Army Research Office under grants W911NF-21-1-0048 and W911NF-18-1-0028.

Appendix A: Examples and Constructions

To clarify the development and build intuition about patterns, as we use the term, the following works through several examples in detail. Specifically, we show explicitly how to construct a semigroup G for exact and generalized symmetries, how to construct a semiautomaton presenting G (and thus $\mathcal{X}$), and how this general semiautomaton simplifies to the future cover, whose irreducible component is the minimal presentation $\mathcal{P}(\mathcal{X})$.

1. Exact Symmetry Shifts

The simplest class of exact symmetry strings to characterize as sofic shifts are the *k-clock shifts* for which $\mathcal{A} = \{0, 1, \ldots, (k-1)\}$ and $b = 01\cdots(k-1)$. For example, points in the 3-clock shift are of the form $\cdots 012012012012\cdots$ with $b = 012$. Intuitively, this is the simplest case since each $a \in \mathcal{A}$ fully specifies the period of the translation symmetry. Recall that this corresponds to the internal states of $\mathcal{P}(\mathcal{X})$, which are the equivalence classes $[\cdot]_F$. For k-clock sequences each internal state (excluding the forbidden state), and thus equivalence class, is represented by a generator $a \in \mathcal{A}$. This fully specifies $\mathcal{X}$ since there are no transient states. Each $a \in \mathcal{A}$ is *synchronizing*. In terms of the defining semigroup of $\mathcal{X}$ for k-clock shift, G is consists solely of $\mathcal{A}$ plus the absorbing element e. The 3-clock shift is given by $G = \{0, 1, 2, e\}$ with $01 = 1$, $12 = 2$, $20 = 0$, $0^2 = 1^2 = 2^2 = e$, and $02 = 10 = 21 = e$.

For more general exact symmetry shifts there are additional elements in G beyond the generators $\mathcal{A} \cup \{e\}$. And, the future cover will have transient states; i.e., there are additional equivalence class $[\cdot]_F$ beyond those in the minimal presentation $\mathcal{P}(\mathcal{X})$. To illustrate, consider the shift $\mathcal{X}$ with $b = 001$ and points of the form $\cdots 001001001\cdots$.

A simple construction of a finite G for an exact symmetry shift $\mathcal{X}$ is as follows. Start by constructing the asymptotic recurrent component of the presenting semiautomaton—this is $\mathcal{P}(\mathcal{X})$. We already showed the states of $\mathcal{P}(\mathcal{X})$ are the equivalence classes $[b]_F$, $[bb_1]_F$, $[bb_1b_2]_F$, and so on. These equivalence classes can generally be represented by the allowed words of length $p-1$ so that concatenation with a generator gives a (shift of) the tiling block b. This represents shifts of windows of length $p-1$ on points $x \in \mathcal{X}$.

For our example with $b = 001$, these words are 00, 01, and 10. The word 11 is not included because it is forbidden in $\mathcal{X}$, and so $11 = 1^2 = e$ in G. If our $p-1$ window is on 00 in $\mathcal{X}$, then a unit shift reveals the generator 1 and the window now shows 01. This gives the production rule $00 \cdot 1 = 001 = 01$ in G. In the semiautomaton presentation there are states $\xi_{00} = [00]_F$ and $\xi_{01} = [01]_F$ with $M_1(\xi_{00}) = \xi_{01}$. If we again shift the window on 01, we reveal the generator 0 and the window now shows 10, giving the production rule $010 = 10$. To complete the cycle we have $100 = 00$, giving the finite closure for G. The rest of the elements in G can be filled in with the free semigroup: e.g., $0 \cdot 1 = 01$. Therefore, we can give a finite G for our example as $G = \{0, 1, e, 00, 01, 10\}$ with production rules $001 = 01, 010 = 10, 100 = 00$ and $1^2 = 0^3 = 101 = e$. The semiautomaton presentation for this G is shown in Fig. 8 (a). Teal colored states are transient, orange is the absorbing forbidden state ξ_e, and black states are the recurrent component (excluding ξ_e). Again, self-loops on ξ_e are omitted for visual clarity.

While this straightforward construct always produces a finite semigroup G for a given exact symmetry shift $\mathcal{X}$, it is not necessarily a minimal semigroup. (Or, equivalently, not a minimal presenting semiautomaton.) The future cover equivalence relation exploits additional structure in $\mathcal{X}$ to give a minimal description and may thus reduce and simplify the straightforward G and its presenting automaton. The k-clock shifts are the extreme example, since we only need the generators as the elements of G since each $a \in \mathcal{A}$ is synchronizing. In the $b = 001$ example, while 0 is not synchronizing, 1 is. Applying the future cover equivalence relations exploits this to simplify G and its presenting semiautomaton. From visual inspection of Fig. 8 (a)'s semiautomaton we see that ξ_1 and ξ_{01} are equivalent, since their transitions lead to the same states with the same labeled edges. Applying the future cover equivalence relation gives the future cover, with its reduced semiautomaton presentation shown in Fig. 8 (b). The simplified semigroup G of the future cover is given as $G = \{0, 1, e, 00, 10\}$ with $01 = 1$ and $100 = 00$, and the same production rules for forbidden words.

Since we designed our straightforward construction around the asymptotic cycle of the translation symmetry, the recurrent components in both cases are isomor-

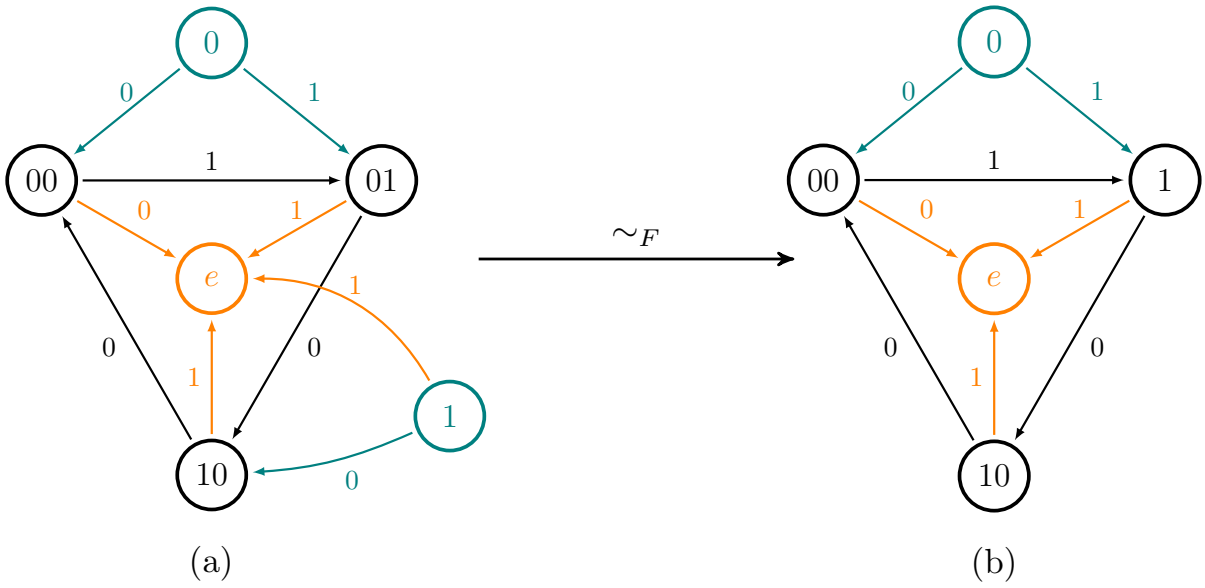


FIG. 8. Semiautomaton presentations for straightforward semigroup construction (a) and its simplification under the future cover equivalence relation (b).

phic and represent the canonical presentation $\mathcal{P}(\mathcal{X})$ that captures the symmetry algebra.

2. General Pattern: the Even Shift

To contrast with exact symmetry shifts, we now go through the well-studied Even Shift [29, 31]. Recall that the Even Shift is the set of sequences that have even-length blocks of 1s bounded by 0s. Thus, it is defined by the set of irreducible forbidden words $\mathcal{F} = \{010, 01^30, 01^50, \dots\}$. Since $\mathcal{F}$ is not finite, the Even Shift is not of finite type—it is *strictly sofic*. However, since it is sofic it can still be finitely defined in terms of a finite semigroup G. Following Ref. [31] we use $G = \{0, 1, e, 01, 10, 11, 101\}$ with production rules $0^2 = 0$, $1^3 = 1$, $01^2 = 1^20 = 0$ and $010 = e$.

The production rule $010 = e$ represents the shortest forbidden word and the other rules allow for all the other forbidden rules to reduce to 010. For example, we use $1^3 = 1$ to reduce 01110 to 010 which then maps to the absorbing element e. We can see that while there is a countably-infinite number of forbidden words in the Even Shift, there is structure in these forbidden words that can be captured in a finite semigroup G.

As with our explicit symmetry example $b = 001$, the presenting semiautomaton for the Even Process using the G above is not minimal. Shown in Fig. 9, we see that there are two recurrent components $\mathcal{P}_A(\mathcal{X})$ and $\mathcal{P}_B(\mathcal{X})$. While these components correspond to different elements in G, we can see that they are again isomorphic and so collapse together under $\sim_F$. The resulting recurrent component is the canonical machine presentation $\mathcal{P}(\mathcal{X})$, shown in Fig. 1 (c).

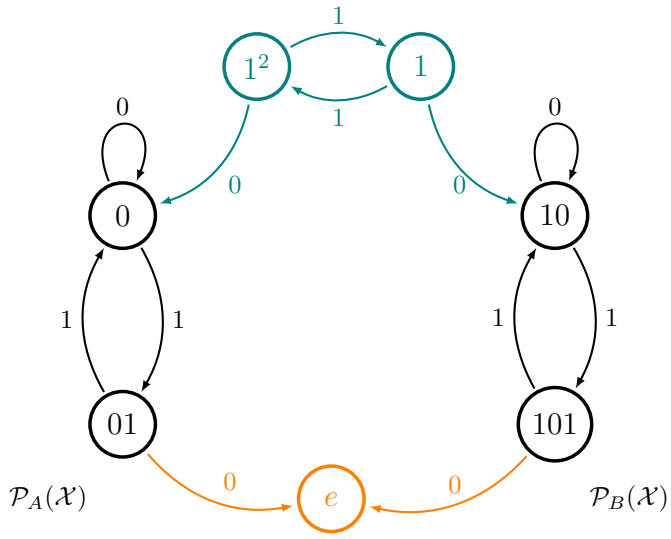


FIG. 9. Presenting semiautomaton for the Even Shift using $G = \{0, 1, e, 01, 10, 11, 101\}$ and $0^2 = 0$, $1^3 = 1$, $01^2 = 1^20 = 0$, $010 = e$. The two components $\mathcal{P}_A(\mathcal{X})$ and $\mathcal{P}_B(\mathcal{X})$ are isomorphic and thus collapse together under $\sim_F$. The recurrent component is the canonical machine presentation $\mathcal{P}(\mathcal{X})$, shown in Fig. 1 (c).

Though we discussed and described three forms of presenting semiautomata, we want to emphasize the importance of the canonical presentation $\mathcal{P}(\mathcal{X})$ as the mathematical representation of pattern as generalized symmetry. There may be many semigroups G that describe a given sofic shift $\mathcal{X}$, and so there are many different semiautomata that can present $\mathcal{X}$. However, for irreducible

sofic shifts the recurrent components of all such G will be isomorphic [31], which is why the future cover is guaranteed to have a single recurrent component. Therefore the future cover *is* a unique representation of the structure of $\mathcal{X}$. Patterns and the symmetries they generalize are asymptotic properties and so the transients of the future cover are not of interest for our current purposes. This is why the canonical presentation $\mathcal{P}(\mathcal{X})$—the recurrent component of the future cover—is the unique mathematical representation of patterns.

Recall that the symmetry group of a translation symmetry is captured by the canonical presentation, whose semiautomaton is a circle graph, neglecting the absorbing state and its transitions. Similarly, we can see the canonical presentation of the Even Shift captures the essential details of the pattern. From inspecting its edge-labeled graph, shown in Fig. 1 (c), we can see that an arbitrary number of 0s are allowed from state ξ_A, but once a 1 occurs it must be followed by another 1, ensuring the pattern of an even number of 1s bounded by 0s. This also highlights the partial regularity that motivated our definition of patterns. Starting in state ξ_A there is a coin flip, either 0 or 1 may occur. If it is a 0, repeat. However, if it is a 1, there is now additional regularity and structure that enforces another 1 to follow.

Appendix B: Distinct Statistical Patterns Supported on the Same Sofic Shift Topological Pattern

The following illustrates how statistical patterns, in the form of sofic measures, are additional structure on top of topological patterns. In particular, there can be complex statistical structure supported on simple topological structure. The topological pattern in this example is the full-2 shift—the set of all binary strings. As described above, the full-2 shift represents a "null" pattern, in the sense that there is no predictability leveraged from knowing the pattern. This is captured quantitatively by the single-state canonical machine presentation $\mathcal{P}(\mathcal{X})$. Its memory—the log of the number of states—is zero.

The simplest statistical patterns supported on the full-2 shift is given by assigning IID single-symbol probabilities, e.g., $\Pr(1) = 0.3$ and $\Pr(0) = 0.7$. The simplicity of this statistical pattern is also captured by a single-state ϵ-machine, where the above probabilities are assigned to the correspond transitions from the state back to itself on the given symbol. In this example it is easy to see the statistical pattern is supported on the full-2 shift. Simply remove the probabilities from the single-state transitions and the single-state canonical machine presentation of the full-2 shift is recovered. Since the ϵ-machine (statistical) and canonical machine presentation (topological) have the same number of the states, the statistical and topological patterns they capture can be thought of as comparable.

However, consider the ϵ-machine shown in Fig. 10. Its two states signify a more complex statistical pattern with

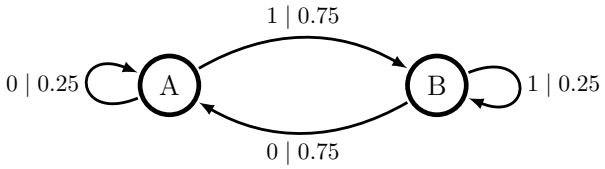


FIG. 10. Two state stochastic ϵ-machine presentation for a statistical pattern supported on the full-2 shift.

nonzero memory. From the machine diagram, we can see that the single-symbol probabilities depend on which of the two internal causal states A and B the process is in. For example, the probability of seeing a 1 is 0.75 if in causal state A and 0.25 if in B. Note also that the symbol 0 always leads to causal state A and 1 always leads to B. Thus, single symbols are *synchronizing* in this example. This stochastic process is an order-1 Markov process. The simple one-state case above is an IID (order-0 Markov) process, by contrast.

It is perhaps clear from Fig. 10's machine diagram that this statistical pattern is supported on the full-2 shift, since both symbols have positive probability from each of the two states. However, it is instructive to show this using the symbol-labeled transition matrices. Recall that T^a_{ij} gives the probability of transitioning from state ξ_i to ξ_j on the symbol $a \in \mathcal{A}$. For Fig. 10's ϵ-machine we have:

$$T^0 = \begin{pmatrix} 0.25 & 0.0 \\ 0.75 & 0.0 \end{pmatrix} ,$$

and

$$T^1 = \begin{pmatrix} 0.0 & 0.75 \\ 0.0 & 0.25 \end{pmatrix} ,$$

where state A is given index 1 and B is index 2, so that T^a_{12} is the probability of transitioning from A to B.

Matrix representations M^a of the topological transition maps are given by setting non-zero elements of T^a to unity. In this example, we have:

$$M^0 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} ,$$

and

$$M^1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} .$$

Recall that $M^a_{ij} = 1$ signifies a transition from internal state i to state j is allowed on the symbol a and $M^a_{ij} = 0$ is a forbidden transition that produces a forbidden word.

To see that these topological transitions correspond to the full-2 shift, note that the action of M^a on both internal states yields the same output state, for both M^0 and M^1. That is, for both states, call them also $A = (1\ 0)$ and $B = (0\ 1)$, a transition on a 0 goes to state A and a transition on a 1 goes to B. For example:

$$(0\ 1) \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = (1\ 0) ,$$

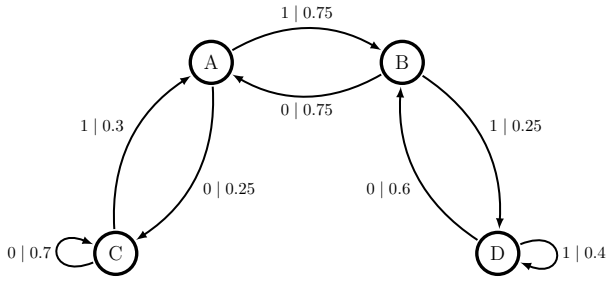


FIG. 11. Four-state stochastic ϵ-machine presentation for a statistical pattern supported on the full-2 shift.

and

$$(1\ 0)\begin{pmatrix}1 & 0\\ 1 & 0\end{pmatrix} = (1\ 0)\ .$$

Thus, as described above, the two states are topologically equivalent $A \sim_F B$ and so reduce to the single-state canonical machine presentation of the full-2 shift.

Finally, as shown in Fig. 11 this construction of probabilistically-distinct causal states supported on the full-2 shift can be extended to ϵ-machine with a larger number of states. Clearly, the construction can be extended indefinitely with an arbitrary number of causal states ξ_i such that $\Pr(0|\xi_i) = p$ and $\Pr(1|\xi_i) = 1 - p$, since there can be uncountably-many p. This shows that there can be arbitrarily complex statistical patterns supported on simple topological patterns.

Appendix C: Cellular automata

A one-dimensional *cellular automaton* or CA $(\mathcal{A}^{\mathcal{L}}, \Phi)$ consists of a spatial lattice $\mathcal{L} = \mathbb{Z}$ whose *sites* take values from a finite *alphabet* $\mathcal{A}$. A CA *state* $x \in \mathcal{A}^{\mathbb{Z}}$ is the configuration of all site values $x^r \in \mathcal{A}$ on the lattice. (For states x, subscripts denote time; superscripts sites.) CA states evolve in discrete time steps according to the *global evolution* $\Phi : \mathcal{A}^{\mathbb{Z}} \to \mathcal{X} \subseteq \mathcal{A}^{\mathbb{Z}}$, where:

$$x_{t+1} = \Phi(x_t)\ .$$

Φ is implemented through parallel, synchronous application of a *local update rule* ϕ that evolves individual sites x_t^r based on their radius R *neighborhoods* $\eta(x^r) = \{x^{r'} \ : \ |r - r'| < R\}$:

$$x_{t+1}^r = \phi\big(\ \eta(x_t^r)\ \big)\ .$$

Stacking the states in a CA *orbit* $x_{0:t} = \{x_0, x_1, \ldots, x_{t-1}\}$ in time-order produces a *spacetime field* $\mathbf{x}_{0:t} \in \mathcal{A}^{\mathbb{Z}\otimes\mathbb{Z}}$. Visualizing CA orbits as spacetime fields reveals the fascinating patterns and localized structures that CAs produce and how the patterns and structures evolve and interact over time.

1. Elementary cellular automata

The parameters $(\mathcal{A}, R)$ define a CA *class*. One simple but nontrivial class is that of the so-called *elementary* cellular automata (ECAs) [66] with a binary local alphabet $\mathcal{A} = \{0, 1\}$ and radius $R = 1$ local-interaction neighborhood $\eta(x_t^r) = x_t^{r-1}x_t^r x_t^{r+1}$. Due to their definitional simplicity and wide study, we mostly explore ECAs in our examples.

A local update rule ϕ is generally specified through a *lookup table* that enumerates all possible neighborhood configurations η and their outputs $\phi(\eta)$. The lookup table for ECAs is given as:

η	$O_\eta = \phi(\eta)$
1 1 1	O_7
1 1 0	O_6
1 0 1	O_5
1 0 0	O_4
0 1 1	O_3
0 1 0	O_2
0 0 1	O_1
0 0 0	O_0

,

where each output $O_\eta = \phi(\eta) \in \mathcal{A}$ and the ηs are listed in lexicographical order. There are $2^8 = 256$ possible ECA lookup tables, as specified by the possible strings of output bits: $O_7O_6O_5O_4O_3O_2O_1O_0$. A specific ECA lookup table is often referred to as an ECA *rule* with a *rule number* given as the binary integer $o_7o_6o_5o_4o_3o_2o_1o_0 \in [0, 255]$. For example, ECA 172's lookup table has output bit string 10101100.

The n^{th}-order lookup table ϕ^n maps the radius $n \cdot R$ neighborhood of a site to that site's value n time steps in the future. Said another way, a spacetime site x_{t+n}^r is completely determined by the radius $n \cdot R$ neighborhood n time-steps in the past according to:

$$x_{t+n}^r = \phi^n\big(\eta^n(x_t^r)\big)\ .$$

The *depth-n past lightcone* is the collection of all ϕ^ℓ for $1 \leq \ell \leq n$, plus the present site value itself (i.e., "depth-0").

Appendix D: ECA Domain Sofic Shifts

The generalization of invariant sets from low-dimensional dynamical systems to high-dimensional cellular automata are known as *domains* [55–57]. These are sets of spatial configurations that are invariant under the global CA dynamic Φ. Spatial configurations of one-dimensional cellular automata are strings of symbols from an alphabet $\mathcal{A}$. And, thus, shift spaces are natural choices for sets of spatial configurations. In fact, the Curtis-Hedlund-Lyndon theorem shows that a CA's dynamic—a sliding-block code—naturally induces shift-invariance in the set of images under Φ [67]. Said another

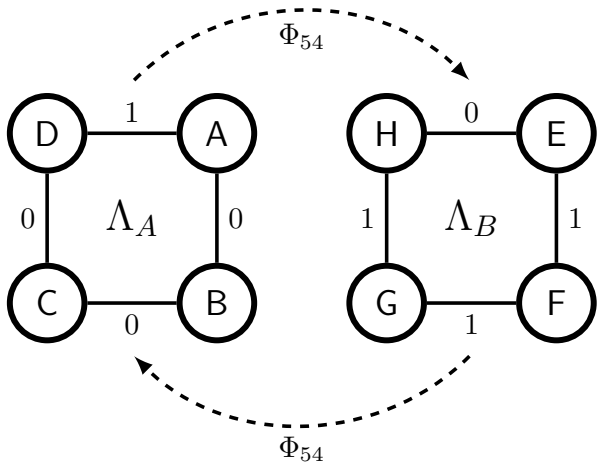


FIG. 12. Presenting semiautomaton for the domain of ECA Rule 54.

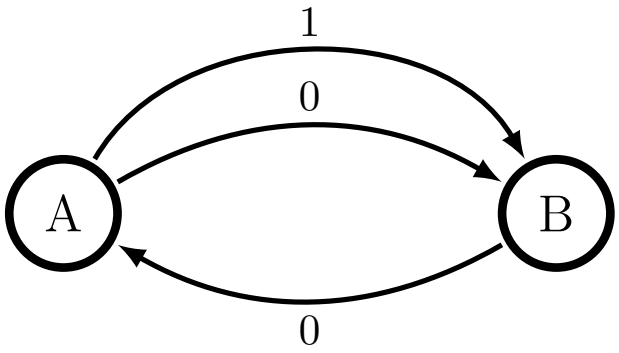


FIG. 13. Presenting semiautomaton for the domain of ECA Rule 18.

way, the dynamic of a CA on a shift space maps to a shift space.

For a CA Φ, a *domain* $\Lambda = \{\Lambda_1, \Lambda_2, \ldots, \Lambda_p\}$ is a set of irreducible sofic shifts such that $\Phi(\Lambda_i) \in \Lambda$. That is, if x is a spatial configuration that is a point in one of the sofic shifts $\Lambda_i \in \Lambda$, then the image $\Phi(x)$ is also a point in one of the sofic shifts in Λ. As irreducible sofic shifts, domains may possess patterns that are exact symmetries, partial symmetries, hidden symmetries, or general patterns with no symmetries. Empirically, the spacetime fields produced from the evolution of ECA domains possess the same type of pattern as their invariant sofic shifts. The generalized spacetime symmetries are revealed by the local causal states, as shown above in Fig. 5. Here, we provide the machine presentations for the domain sofic shifts used in each case.

The exact symmetry spacetime field is given by the evolution of ECA Rule 54's domain, whose machine presentation is shown in Fig. 12. As can be seen, there are two phase Λ_A and Λ_B that are cycled between under Φ_{54}. Each phase has an exact symmetry, with Λ_A tiled by blocks of 0001 and Λ_B tiled by blocks of 1110. As seen in Fig. 5 (a), this creates an exact symmetry spacetime field that is period-4 in both time and space.

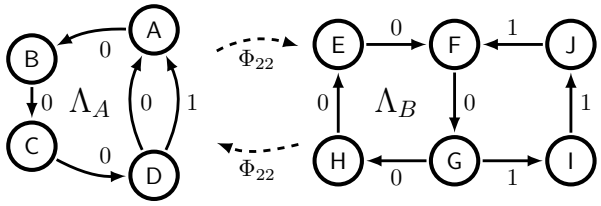


FIG. 14. Presenting semiautomaton for the domain of ECA Rule 22.

The partial symmetry spacetime field possesses a stochastic symmetry and is given by the evolution of ECA Rule 18's domain. Its machine presentation is shown in Fig. 13. This is a single stochastic symmetry sofic shift with a period-2 tiling of 0Σ, where again Σ is a wildcard that can be either 0 or 1. Spacetime fields generated by the domain of Rule 18, as shown in Fig. 5 (b), form a subshift of the 0-Σ checkerboard spacetime shift space. Both are discussed in detail above in the stochastic symmetry Case Study of Section IV.

Finally, the hidden symmetry spacetime field in Fig. 5 (c) is generated by the hidden symmetry domain sofic shift of ECA Rule 22. The presenting automaton is shown in Fig. 14. Like Rule 54's domain, the domain of ECA Rule 22 comes in two phases. As can be seen, phase Λ_A is actually a stochastic symmetry, as each state in $\mathcal{P}(\Lambda_A)$ returns to itself after four translations. However, this is not the case for phase Λ_B, which is a hidden symmetry sofic shift. This is because only states F and G of $\mathcal{P}(\Lambda_B)$ return to themselves after four translations. As drawn, one can imagine "folding" $\mathcal{P}(\Lambda_B)$ up onto itself to create a period-4 symmetry in the states. Said another way, one can think of a "length-3 wildcard" that produces either three 0s or three 1s. The Λ_B sofic shift is then given by tilings of this length-3 wildcard followed by a fixed 0. This is the nature of hidden symmetries. Blocks of three 0s and three 1s are forced to occur, followed by a fixed 0, but the sequence of these blocks is not constrained. For example, the string of all 0s and the string of all 1s are both in Λ_B.

[1] M. C. Cross and P. C. Hohenberg, Pattern formation outside of equilibrium, Rev. Mod. Phys. **65**, 851 (1993).
[2] P. Ball, *The Self-Made Tapestry: Pattern Formation in Nature* (Oxford University Press, New York, 1999).
[3] M. Cross and H. Greenside, *Pattern Formation and Dynamics in Nonequilibrium Systems* (Cambridge Univer-

sity Press, Cambridge, United Kingdom, 2009).
[4] H. Bernard, Les tourbillons cellulaires dans une nappe liquide [the cellular vortices in a liquid layer], Rev. Gén. Sci. Pure Appl. **11**, 1261 (1900).
[5] M. Paul, K.-H. Chiam, M. Cross, P. Fischer, and H. Greenside, Pattern formation and dynamics in Rayleigh–Bénard convection: Numerical simulations of experimentally realistic geometries, Physica D: Nonlinear Phenomena **184**, 114 (2003).
[6] A. M. Zhabotinsky, Periodical oxidation of malonic acid in solution (a study of the Belousov reaction kinetics), Biofizika **9**, 306 (1964).
[7] I. Epstein and J. Pojman, *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos* (Oxford University Press, 1998).
[8] P. Holmes, J. Lumley, G. Berkooz, and C. Rowley, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* (Cambridge University Press, Cambridge, United Kingdom, 2012).
[9] G. Haller, Lagrangian coherent structures, Ann. Rev. Fluid Mech. **47**, 137 (2015).
[10] P. S. Marcus, Numerical simulation of Jupiter's Great Red Spot, Nature **331**, 693 (1988).
[11] J. Sommeria, S. D. Meyers, and H. L. Swinney, Laboratory simulation of Jupiter's Great Red Spot, Nature **331**, 689 (1988).
[12] W.-K. Tung, *Group Theory in Physics: an Introduction to Symmetry Principles, Group Representations, and Special Functions in Classical and Quantum Physics* (World Scientific Publishing Co Inc, Philidelphia, 1985).
[13] A. Rupe and J. P. Crutchfield, Local causal states and discrete coherent structures, Chaos **28**, 1 (2018).
[14] A. Rupe, N. Kumar, V. Epifanov, K. Kashinath, O. Pavlyk, F. Schlimbach, M. Patwary, S. Maidanov, V. Lee, Prabhat, and J. P. Crutchfield, Disco: Physics-based unsupervised discovery of coherent structures in spatiotemporal systems, in *2019 IEEE/ACM Workshop on Machine Learning in High Performance Computing Environments (MLHPC)* (IEEE, 2019) pp. 75–87, arXiv:1909.11822 [physics.comp-ph].
[15] K. Schmidt, *Dynamical systems of algebraic origin* (Springer Science & Business Media, 1995).
[16] K. Lindgren, C. Moore, and M. Nordahl, Complexity of two-dimensional patterns, J. Stat. Phys. **91**, 909 (1998).
[17] D. Lind, Multi-dimensional symbolic dynamics, in *Symbolic Dynamics and its Applications*, Vol. 60 (American Mathematical Society, 2004) pp. 61–80.
[18] C. R. Shalizi and J. P. Crutchfield, Computational mechanics: Pattern and prediction, structure and simplicity, J. Stat. Phys. **104**, 817 (2001).
[19] U. Grenander, *Elements of pattern theory* (JHU Press, 1996).
[20] D. Mumford and A. Desolneux, *Pattern theory: the stochastic analysis of real-world signals* (CRC Press, 2010).
[21] J. M. Greenberg and S. P. Hastings, Spatial patterns for discrete models of diffusion in excitable media, SIAM Journal on Applied Mathematics **34**, 515 (1978).
[22] J. P. Crutchfield and D. P. Feldman, Statistical complexity of simple one-dimensional spin systems, Phys. Rev. E **55**, R1239 (1997).
[23] D. P. Feldman, *Computational Mechanics of Classical Spin Systems*, Ph.D. thesis, University of California, Davis (1998), Published by University Microfilms Intl, Ann Arbor, Michigan.
[24] D. P. Feldman and J. P. Crutchfield, Structural information in two-dimensional patterns: Entropy convergence and excess entropy, Phys. Rev. E **67**, 051103 (2003).
[25] M. D. Robinson, D. P. Feldman, and S. R. McKay, Local entropy and structure in a two-dimensional frustrated system, Chaos **21**, 037114 (2011).
[26] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding* (Cambridge University Press, New York, 1995).
[27] S. Smale, Differentiable dynamical systems, Bull. Amer. Math. Soc. **73**, 797 (1967).
[28] W. Parry, Intrinsic Markov chains, Trans. Amer. Math. Soc. **112**, 55 (1964).
[29] B. Weiss, Subshifts of finite type and sofic systems, Monastsh. Math. **77**, 462 (1973).
[30] A. Ginzburg, *ACM Monograph Series* (Academic Press, 1968).
[31] B. Kitchens and S. Tuncel, Semi-groups and graphs, Israel. J. Math. **53**, 231 (1986).
[32] W. Krieger, On sofic systems I, Israel J. Math. **48**, 305 (1984).
[33] R. Fischer, Sofic systems and graphs, Monastsh. Math **80**, 179 (1975).
[34] R. Fischer, Graphs and symbolic dynamics, Coll. Math. Soc. Janos Bolyai, 16. Topics in Information Theory (1975).
[35] R. B. Ash, *Information Theory* (John Wiley and Sons, New York, 1965).
[36] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, 3rd ed. (Prentice-Hall, New York, 2006).
[37] M. Marques-Pita and L. Rocha, Schema redescription in cellular automata: Revisiting emergence in complex systems, in *Artificial Life (ALIFE), 2011 IEEE Symposium on* (IEEE, 2011) pp. 233–240.
[38] K. Krohn and J. Rhodes, Algebraic theory of machines. i. prime decomposition theorem for finite semigroups and machines, Trans. Am. Math. Soc. **116**, 450 (1965).
[39] J. Rhodes, *Applications of Automata Theory and Algebra via the Mathematical Theory of Complexity to Biology, Physics, Psychology, Philosophy, Games, and Codes* (University of California, Berkeley, California, 1971) c. Nehaniv, editor, World Scientific Publishing Company, Singapore (2009).
[40] M. Boyle and K. Petersen, Hidden Markov processes in the context of symbolic dynamics, in *Entropy of Hidden Markov Processes and Connections to Dynamical Systems*, edited by B. Marcus, K. Petersen, and T. Weissman (Cambridge University Press, 2011).
[41] R. M. Dudley, *Real Analysis and Probability*, first 1989 ed. (Cambridge University Pressl, Cambridge, 2002).
[42] H. Furstenberg, Stationary processes and prediction theory, Ann. Math. Studies **44** (1960).
[43] J. P. Crutchfield and K. Young, Inferring statistical complexity, Phys. Rev. Let. **63**, 105 (1989).
[44] P. Riechers and J. P. Crutchfield, Spectral simplicity of apparent complexity, Part I: The nondiagonalizable metadynamics of prediction, Chaos **28**, 033115 (2018).
[45] P. Riechers and J. P. Crutchfield, Spectral simplicity of apparent complexity, Part II: Exact complexities and complexity spectra, Chaos **28**, 033116 (2018).
[46] J. P. Crutchfield and D. P. Feldman, Regularities unseen, randomness observed: Levels of entropy conver-

gence, CHAOS **13**, 25 (2003).
[47] D. P. Feldman, C. S. McTague, and J. P. Crutchfield, The organization of intrinsic computation: Complexity-entropy diagrams and the diversity of natural information processing, CHAOS **18**, 043106 (2008).
[48] J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney, Time's barbed arrow: Irreversibility, crypticity, and stored information, Phys. Rev. Lett. **103**, 094101 (2009).
[49] J. P. Crutchfield, C. J. Ellison, J. R. Mahoney, and R. G. James, Synchronization and control in intrinsic and designed computation: An information-theoretic analysis of competing models of stochastic computation, CHAOS **20**, 037105 (2010).
[50] R. G. James, C. J. Ellison, and J. P. Crutchfield, Anatomy of a bit: Information in a time series observation, CHAOS **21**, 037109 (2011).
[51] J. P. Crutchfield, P. Riechers, and C. J. Ellison, Exact complexity: Spectral decomposition of intrinsic computation, Phys. Lett. A **380**, 998 (2015).
[52] P. M. Ara, R. G. James, and J. P. Crutchfield, The elusive present: Hidden past and future dependence and why we build models, Phys. Rev. E **93**, 022143 (2016).
[53] C. Shalizi, Optimal nonlinear prediction of random fields on networks, Discrete Mathematics & Theoretical Computer Science (2003).
[54] A. Rupe and J. P. Crutchfield, Spacetime symmetries, invariant sets, and additive subdynamics of cellular automata, arXiv preprint arXiv:1812.11597 (2018).
[55] J. E. Hanson and J. P. Crutchfield, The attractor-basin portrait of a cellular automaton, J. Stat. Phys. **66**, 1415 (1992).
[56] J. P. Crutchfield and J. E. Hanson, Turbulent pattern bases for cellular automata, Physica D **69**, 279 (1993).
[57] J. E. Hanson and J. P. Crutchfield, Computational mechanics of cellular automata: An example, Physica D **103**, 169 (1997).
[58] A. Rupe, Rule 90 patterns jupyter notbook, `https://github.com/adamrupe/ca_patterns/blob/main/notebooks/rule90_patterns.ipynb` (2022).
[59] A. Rupe, Rule 18 domain patterns jupyter notbook, `https://github.com/adamrupe/ca_patterns/blob/main/notebooks/domain18_patterns.ipynb` (2022).
[60] A. Rupe, General ca patterns jupyter notbook, `https://github.com/adamrupe/ca_patterns/blob/main/notebooks/general_patterns.ipynb` (2022).
[61] D. P. Varn and J. P. Crutchfield, Chaotic crystallography: How the physics of information reveals structural order in materials, Curr. Opin. Chem. Eng. **7**, 47 (2015).
[62] A. B. Boyd, D. Mandal, and J. P. Crutchfield, Identifying functional thermodynamics in autonomous Maxwellian ratchets, New J. Physics **18**, 023049 (2016).
[63] A. J. Palmer, C. W. Fairall, and W. A. Brewer, Complexity in the atmosphere, IEEE Trans. Geosci. Remote Sensing **38**, 2056 (2000).
[64] S. E. Marzen and J. P. Crutchfield, Optimized bacteria are environmental prediction engines, Phys. Rev. E **98**, 012408 (2018).
[65] A. E. Venegas-Li, A. M. Jurgens, and J. P. Crutchfield, Measurement-induced randomness and structure in controlled qubit processes, Phys. Rev. E **102**, 040102 (2020).
[66] S. Wolfram, Statistical mechanics of cellular automata, Rev. Mod. Phys. **55**, 601 (1983).
[67] G. A. Hedlund, Endomorphisms and automorphisms of the shift dynamical system, Theory of Computing Systems **3**, 320 (1969).