Markov modeling via ordinal partitions: An alternative paradigm for network-based time-series analysis

Konstantinos Sakellariou [1,2,*] Thomas Stemler,[1,†] and Michael Small[1,‡]

[1]*School of Mathematics & Statistics, The University of Western Australia, Crawley WA 6009, Australia*
[2]*Nodes & Links Ltd, Leof. Athalassas 176, Strovolos, Nicosia, 2025, Cyprus*

Mapping time series to complex networks to analyze observables has recently become popular, both at the theoretical and the practitioner's level. The intent is to use network metrics to characterize the dynamics of the underlying system. Applications cover a wide range of problems, from geoscientific measurements to biomedical data and financial time series. It has been observed that different dynamics can produce networks with distinct topological characteristics under a variety of time-series-to-network transforms that have been proposed in the literature. The direct connection, however, remains unclear. Here, we investigate a network transform based on computing statistics of ordinal permutations in short subsequences of the time series, the so-called *ordinal partition network*. We propose a Markovian framework that allows the interpretation of the network using ergodic-theoretic ideas and demonstrate, via numerical experiments on an ensemble of time series, that this viewpoint renders this technique especially well-suited to nonlinear chaotic signals. The aim is to test the mapping's faithfulness as a representation of the dynamics and the extent to which it retains information from the input data. First, we show that generating networks by counting patterns of increasing length is essentially a mechanism for approximating the analog of the Perron-Frobenius operator in a topologically equivalent higher-dimensional space to the original state space. Then, we illustrate a connection between the connectivity patterns of the networks generated by this mapping and indicators of dynamics such as the hierarchy of unstable periodic orbits embedded within a chaotic attractor. The input is a scalar observable and any projection of a multidimensional flow suffices for reconstruction of the essential dynamics. Additionally, we create a detailed guide for parameter tuning. We argue that there is no optimal value of the *pattern length m*, rather it admits a scaling region akin to traditional embedding practice. In contrast, the embedding lag and overlap between successive patterns can be chosen exactly in an optimal way. Our analysis illustrates the potential of this transform as a complementary toolkit to traditional time-series methods.

I. INTRODUCTION

Data analysis methods that employ concepts and measures from network science [1–8] have been increasingly attracting the interest of nonlinear time-series-analysis practitioners over the past decade. Applications include human electrocardiograms (ECGs) [2,9,10], time series of earthquake epicentre locations [11], marine paleoclimate data [12], surface air temperature measurements [13], gas-liquid phase transitions [14], coastal hurricane activity [15], foreign exchange rates [16], lake sediment records [17], global equity index time series [18], and diode resonator circuit data [19]. The chief aim is the detection of dynamical characteristics given scalar time series by encoding information into an abstract structure and computing various statistics thereof. A complex network representation provides immense potential towards this end due to the variety of existing measures and their computational simplicity. Additionally, it provides a natural framework for defining the notion of connectivity between nonlinearly interacting components and studying characteristics of topological nature. The pertinent question, however, is whether network representations can indeed capture fundamentally different system dynamics and offer complementary insight to existing techniques.

Various studies have demonstrated a proof of concept by using networks to differentiate dynamical behavior and mapping features of time series to structures exhibiting distinct topological properties [20,21]. However, it remains unclear how the resulting networks relate to the underlying dynamics. Here, we investigate a mapping based on symbolic dynamics, in particular ordinal analysis [22–25], and attempt to provide a solid theoretical foundation to the class of ordinal network transforms [7,9,10,19,26–29]. We reformulate the technique first introduced by Small [7] via a rigorous Markovian framework that allows the approximation of deterministic dynamics through examination of statistical properties of a stochastic model [30]. The key concept is the description of system-wide dynamics through the evolution of sets, rather than characterizing dynamical behavior at the level of trajectories. Traditionally this is achieved by using Ulam's method for discretizing the *Perron-Frobenius* (or *transfer*) operator [31,32]. Instead, we use a modified version defined on ordinal partitions.

_____

*kostas@nodeslinks.com
†thomas.stemler@uwa.edu
‡michael.small@uwa.edu

Existing ordinal network formulations [7,9,10,19,26–29] are not based on a rigorous framework, but heuristic approaches. The network tools employed so far are, therefore, not inspired by a theory, e.g., ergodic. Instead, simple measures used widely in hot areas of network science, e.g., number of nodes, mean number of connections (or its variance), average shortest path length, etc. have constituted the main tools of analysis. We have specifically opted for a rigorous approach to (a) be able to choose network metrics for the purposes of analysis in a targeted, informed manner, (b) explain characteristics of the network topology produced when encoding deterministic chaotic dynamics and (c) provide a link between ordinal network features and the partitioning of the original state space.

Our experiments illustrate the connection between fundamental dynamical features underlying the input time series and the structure of the network. We uncover a relationship between the hierarchy of unstable periodic orbits (UPOs) embedded within chaotic attractors and the cycles found in the network. An inverse mapping of network properties to the original dataset enables accurate localisation of low-order UPOs. With computational efficiency in mind, one of the reasons for promoting this framework is to set the stage for generating quantitative estimates of characteristic dynamical indicators, e.g., Lyapunov exponents [33] and topological entropy [34], using ergodic-theoretic ideas.

This work is also intended to be a guide to the practitioner, highlighting subtle aspects of parameter tuning and shedding light on sampling constraints. All other ordinal network studies recommend parameter selection either based on heuristic arguments or upon inspection of results, without rigorous reasoning or consistent selection among different systems. One of the reasons is that their focus is application-oriented, hence an *ad hoc* choice targeting a specified dataset in mind may be more suitable within such a scope. In contrast, the motivation behind this study is to assess the extent to which ordinal network representations remain faithful to the original dynamics. Our suggestions are the result of sensitivity analysis drawn from numerical experiments on a large variety of test systems, both discrete and continuous. It led to different conclusions with respect to appropriate selection; a main finding being that there is no optimal value for the pattern length, rather it admits a scaling region similar to the embedding dimension in traditional practice. Finally, we advocate using Poincaré maps in the case of continuous flows as constructed networks are less sensitive and more effective for detecting UPOs or estimating invariants [33,34]. In summary, our findings comprise evidence that the proposed mapping can be used to augment traditional analysis toolkits.

II. TIME-SERIES-TO-NETWORK TRANSFORMS

Zhang and Small [2] first utilized complex networks to examine properties of various time series. Mapping pseudoperiodic cycles to nodes on a network, with connections assigned by a similarity criterion, they used statistical measures to examine the topology of the network. Their analysis showed a distinction in topological structure between chaotic and random signals. Several different network transforms have been proposed [1,3–7,14], each best equipped to capture certain

aspects of dynamical behavior—and hence more appropriate for applying to certain systems—while less effective in capturing other aspects [20]. The various methods differ primarily on the nature of the entity for which nodes encode, and therefore the type of information represented by the network: (a) topological characteristics of state space [3,6,35], (b) temporal information [1,7], or (c) geometrical properties of the time series [4]. Network transforms may be broadly classified into three categories [20].

In *proximity networks* nodes correspond to segments in state space. Connectivity is defined by considering mutual proximity of different segments. The most commonly employed method in this class is the recurrence network paradigm [6]. It has proved effective for real-world data [12,17] and, more recently, some theoretical links between classical nonlinear dynamics tools and the network structure have been presented [36]. Inversion of the transform has been proved for a special class of appropriately weighted ϵ-recurrence networks under certain conditions on the given time series [37]—specifically, a uniform distribution of points on the attractor—as well as for equivalent (all other aspects being equal) k-nearest neighbor networks [38]. In *visibility graphs* nodes represent single time series points and connectivity is defined by considerations of convexity constraints [4,18,39]. *Transition networks* comprise formations whereby nodes represent coarse-grained dynamical states and connectivity is defined by temporal succession [1,7,9,19,26,40]. While the choice of network transform is critical, and each method can shed light on different characteristics of dynamical behavior, the concept of *recurrence* is underlying the majority of existing methods [20]. This refers to a system's propensity to revisit certain dynamical states over and over again as time progresses. In mathematical terms, this renders measure-preserving dynamical systems well suited to this type of analysis due to the Poincaré recurrence theorem.

A. Transition networks

The working hypothesis behind almost the entire spectrum of nonlinear time-series-analysis techniques is the existence of a deterministic component in time series, possibly hidden by noise [41]. The framework of a state space may then be established and, under certain technical conditions (endowment with a Borel algebra and a measure-preserving evolution rule), one can define the natural invariant measure, i.e., the probability distribution of all possible states [42]. In practice, most systems are intuitively thought of as evolving continuously in time and capable of obtaining any value from within an interval (continuous state space). This leads to the necessity for discretization, which can be defined by any arbitrary partition of state space [43]. One option is a regular-lattice partition into quantiles or equal-size bins [44,45], i.e., the familiar *histogram* at a specified resolution, which constitutes the most commonly employed choice.

Dynamical states are then labeled in a discrete form by each bin. Transition networks are formed by estimation of transitional probabilities from one state to another by means of frequency statistics computed from time series [1,40,46]. Therefore, in stark contrast to proximity-based approaches, transition networks explicitly involve the temporal order of

observations and, thereby, preserve (to a certain extent) causal relationships contained in the dynamics within connectivity patterns. At the same time, a common drawback of transition networks is that amplitude information, especially finer aspects such as small fluctuations, is discarded. This class of approaches is, consequently, better suited for identifying state space regions that are central to the causal evolution mechanism of the system, e.g., high mixing, stretching and transport of lobe dynamics. Nicolis *et al.* [1], for instance, uncovered a link between local stability properties, as well as global indicators, of the underlying dynamics and the incoming and outgoing connections in the network. Padberg *et al.* [40] followed a different direction using set-oriented methods to identify global transport barriers in continuous flows.

The focal point of this study is a methodology which belongs to the class of transition networks. The foundation is the notion of an *ordinal partition* of state space, dynamics on which can easily be obtained by counting and categorizing short time-series segments. Ordinal partition networks (OPNs) [7] are a model of the coarse-grained dynamics in that connectivity between nodes is defined via temporal succession of symbols that label each category. Sun *et al.* [26] extended the original transform by including amplitude information in the symbolic representation and allowing directed (forward-time causality only) and weighted edges (frequency count of occurring transitions between a pair of symbols). McCullough *et al.* [19] considered a generalized version of this formulation by including a time lag parameter into the network formation algorithm and applied the technique to diode resonator data. Masoller *et al.* [27] applied this OPN variant with a single-step lag to one-dimensional discrete maps and semiconductor laser intensity data to capture changes in a time series. Similarly, Kulp *et al.* [10] applied it to ECG data from patients with a variety of heart conditions. McCullough *et al.* [9] generalized the framework even further by varying the lag to conduct multiscale analysis on ECG data. Zhang *et al.* [28] augmented the OPN construction algorithm to map multivariate time series to a network. Guo *et al.* [29] proposed a different scheme for multivariate time series whereby cross and joint OPNs are computed. Input series are drawn from two coupled systems, where synchronization is often present.

All existing OPN variants share three characteristics. First, the methodology itself, as well as the resulting networks, were not explored in a rigorous fashion. No unifying framework relating network topology to the original dynamics has been proposed so far. Hence, no clear way of estimating dynamical invariants exists due to the lack of a mathematical foundation. Second, in all these studies metrics that belong to two groups were employed to analyze a specific dataset at hand. They comprise either (i) common network science tools (e.g., mean and variance of number of connections, characteristic path length etc.; see Refs. [10,19,26]) or (ii) information-theoretic measures (some form of Shannon entropy on empirical distributions of nodal properties; see Refs. [9,27–29]). Finally, an *ad hoc* selection of parameters is suggested based either on (i) heuristic arguments or (ii) inspection of results, without providing rigorous reasoning or consistent selection among different systems. We propose a framework that (a) allows a natural interpretation of the network as an approximation

to the dynamics, (b) provides an informed, targeted means of choosing useful metrics, (c) enables use of ergodic-theoretic tools and (d) suggests parameter selection according to sensitivity analysis of the most general OPN variant for univariate series applied to a gamut of discrete and continuous systems.

B. Ordinal symbols

Symbolic dynamics is an incredibly powerful tool in the study of dynamical systems. The core idea is the discretization of state space into a finite number of components, each labeled by a different symbol drawn from a finite set, an *alphabet*. Dynamics of infinitely long symbolic sequences, e.g., from an alphabet such as the binary set $\{0,1\}$, can—under certain conditions—be topologically conjugate to the original dynamics [47]. Ordinal symbolization considers a linearly ordered state-space partition based on patterns which encode the relative amplitude of observed values in short consecutive time series segments. Each symbol reflects the rank ordering of sample points within each segment. By recording the relative frequency of occurrence of all symbols, an empirical probability distribution is constructed. The Shannon entropy of this distribution, permutation entropy (PE), quantifies the diversity of possible orderings [22]—referred to as the set of *admissible* ordinal patterns. Given finite data, a practitioner only has access to a sample of this set, the patterns which are in fact realised in the given time series. In the limit of infinitely large pattern length, Bandt *et al.* [23] proved analytically that PE is equal to the Kolmogorov-Sinai (KS) entropy for piecewise monotonic self-interval maps in $\mathbb{R}$, and a similar statement holds for the topological entropy. Amigó *et al.* [24] later generalized this result to discrete deterministic systems defined by ergodic maps on intervals in $\mathbb{R}^n$. This result relies on a modified concept of PE derived from a finite-state stationary stochastic process equipped with an arbitrary order. A different viewpoint was introduced by the standardised approach due to Keller and Sinn [48], whereby multidimensional ordinal patterns, inspired by the original definition, are defined.

Ordinal analysis rapidly attracted the interest of practitioners, especially among investigations of complex biomedical signals, due its computational efficiency. Contributions in the field include several studies on electroencephalograph (EEG) measurements using PE as a criterion for distinction between different brain physiologies; qualitative changes related to epileptic activity in multidimensional time series obtained from different channels [25], detection of the onset of epileptic seizures [49], differentiation of brain states using ordinal pattern distributions [50], detection of deterministic dynamics in epileptic seizures [51], to name a few. Cardiac interbeat-interval biosignals [52,53], climate time series [54], and financial time series [55] have also been analyzed in the same fashion.

Apart from PE, the other main tool employed in applied ordinal analysis is the *count of forbidden patterns* (CFP). Deterministic rules dictate certain constraints in the set of ordinal patterns which can be realised. In particular, such systems are characterized by the existence of a set of forbidden patterns [56] that cannot occur in any time series realisations thereof. This set can be thought of as the collection of regions in

state space where the trajectory is not allowed to transverse. It is an example of *pruning*, a rule which forbids the occurrence of certain subsequences in the symbolic dynamics. However, random systems can admit every possible pattern given sufficiently long time series. This stems from the fact that sampling a stochastic process will generate identically and independently distributed (i.i.d.) values. Mapping the sample data set to a symbolic alphabet of finite size results in a sequence of i.i.d. permutations. The probability that any ordinal pattern is missing approaches zero as the total number of data points grows. This pruning rule, i.e., CFP, has been utilized as a criterion for the detection of a deterministic component in noisy time series data [57].

In the current study we investigate the properties of the graphs associated with the Markovian formulation produced by alphabets which are composed of ordinal symbols of increasing length. We argue that counting ordinal patterns from a time series is essentially equivalent to defining a partition in a higher-dimensional state space wherein the underlying dynamics have been embedded. An adapted form of Ulam's method is then used to provide a definition for ordinal networks computed from time series generated by low-dimensional chaotic ergodic systems (see Fig. 1 for a schematic of the transform). We examine the extent to which this network-based stochastic approximation remains faithful to the true deterministic dynamics.

The outline of the paper is as follows. In Sec. III we present the method. In particular, we show how to extract a symbolic sequence from a time series (Sec. III A), define the notion of an ordinal partition (Sec. III B), introduce the Markovian framework (Sec. IV), redefine the construction of an OPN (Sec. IV A), and last, we discuss ergodicity (Sec. IV B) and irreducibility (Sec. IV C). In Sec. V we present the results of our investigations. Specifically, we first explore the representation of simple dynamics as a proof of concept via the transformation of periodic series (Sec. V A). Sec. V B is devoted to the selection of methodological parameters, as well as to practical considerations presented by real-world data, e.g., finiticity. In Sec. V C we examine the network structures obtained by the numerical experiments conducted on various chaotic discrete-time dynamical systems in one and two dimensions. Furthermore, we adapt the application of OPN transforms to data generated by three-dimensional chaotic continuous-time flows by means of Poincaré maps. Finally, in Sec. V D we illustrate the connection between cycles in OPNs and the dominant (lowest-order) UPOs embedded within a chaotic attractor. This result is used to successfully detect UPOs of order 4 in the Lorenz attractor. Conclusions are discussed in Sec. VI.

III. COUNTING ORDINAL PATTERNS

We consider deterministic dynamical systems defined by the pair $(\mathcal{M}, \phi)$, where the single-valued C^r function $\phi : \mathcal{M} \to \mathcal{M}$ represents the action of a rule for temporal evolution and the compact manifold $\mathcal{M}$ specifies a state space for the system. Points within $\mathcal{M}$, say $\boldsymbol{x}_n \in \mathcal{M} \subseteq \mathbb{R}^d$ for some dimension $d \in \mathbb{N}$, uniquely represent the state of the system at some instant in time. Time is prescribed either in terms of advances in discrete time steps $n \in \mathbb{Z}$ (iterated maps) or as
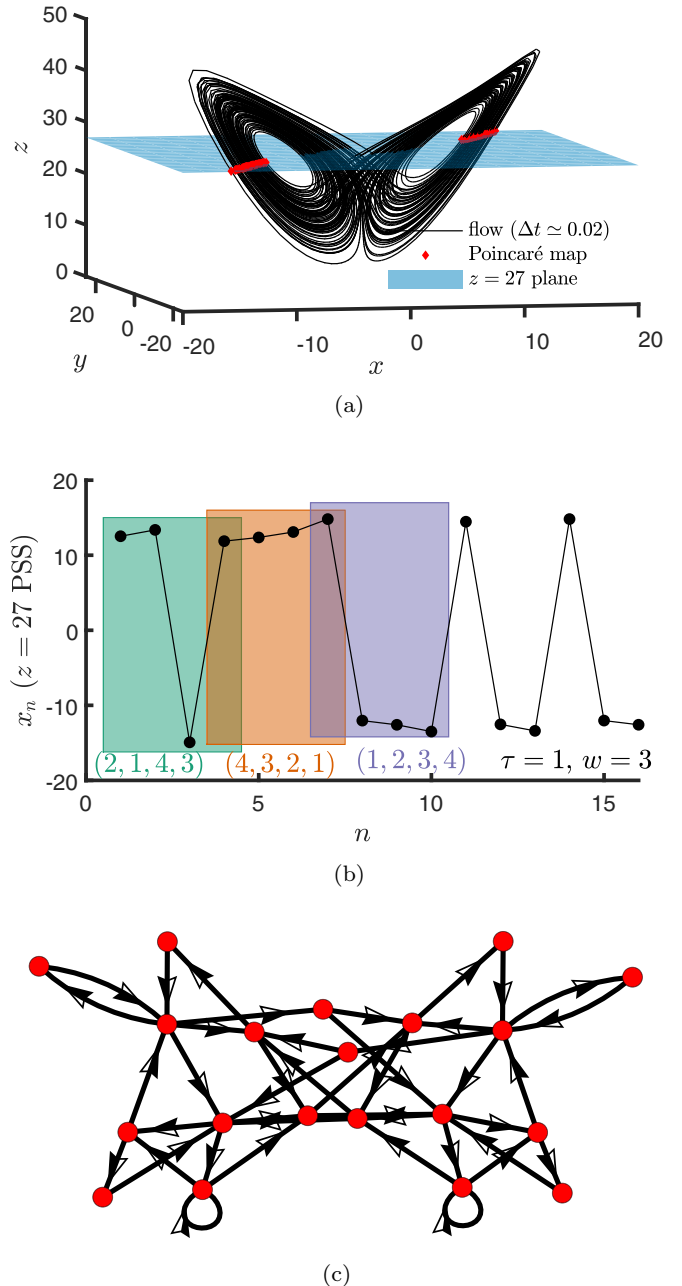

(a)

(b)

(c)

FIG. 1. Schematic of the procedure used to generate an OPN given a scalar measurement from a multidimensional flow. (a) Lorenz attractor. Observables obtained via crossings of a long trajectory ($N = 1, 133, 155$) with the $z = 27$ plane in the $\dot{z} > 0$ direction. (b) Ordinal patterns of length $m = 4$ extracted from the x-component series of the 2D Poincaré map. (c) Two-dimensional spring embedding layout (produced by a built-in MATHEMATICA algorithm) of the OPN that represents coarse-grained dynamics on the Lorenz attractor.

progression in continuous time $t \in \mathbb{R}$, i.e., continuous-time flows generated as solutions to ordinary differential equations. We focus on the former type and examine the latter through the lens of Poincaré maps. Counting patterns via a Poincaré map associated with a flow instead of the numerical approximation obtained by sampling the system at discrete time instants is unorthodox within ordinal symbolization, but

offers the benefit of not having to select the sampling rate (and embedding lag, see Sec. V B 1) parameters.

The equation which defines a discrete-time system of d state variables is given the by the recurrence relation of general form

$$\boldsymbol{x}_{n+1} = \boldsymbol{f}(\boldsymbol{x}_n; \boldsymbol{\theta}), \tag{1}$$

where $\boldsymbol{f} : \mathbb{R}^d \to \mathbb{R}^d$ is a nonlinear vector field which describes the action of the dynamics on states in d-dimensional space, i.e., points defined by the ordered d-tuple $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)$ and a set of parameters $\boldsymbol{\theta}$. Temporal evolution is described in terms of forward-time iterates from a specified initial state given by functional composition, i.e., the *flow* which specifies the solution to Eq. (1) may be represented by

$$\phi(\boldsymbol{x}_0, n; \boldsymbol{\theta}) = \phi_n(\boldsymbol{x}_0; \boldsymbol{\theta}) = \boldsymbol{f}^n(\boldsymbol{x}_0, \boldsymbol{\theta}) \tag{2}$$

for some initial state $\boldsymbol{x}_0$.

A collection of states specified by the solution of Eq. (1) for some initial state $\boldsymbol{x}_0 = \boldsymbol{x}(0)$ is called a *trajectory*. The set of *all* state-space points which lie in the 'past' or 'future' of a specified $\boldsymbol{x}_0$ is referred to as the *orbit* through $\boldsymbol{x}_0$. It is given by the bi-infinite (or infinite for noninvertible maps) sequence

$$\mathcal{O}(\boldsymbol{x}_0) = \{\boldsymbol{x}_n \in \mathbb{R}^d | \boldsymbol{x}_n = \boldsymbol{f}^n(\boldsymbol{x}_0; \boldsymbol{\theta}), \quad n \in \mathbb{Z}\}, \tag{3}$$

where $\boldsymbol{f}^n(\boldsymbol{x})$ denotes the pre-image of the nth iterate of $\boldsymbol{f}$ if n is negative.

Ideally, one wishes to possess full knowledge of the temporal evolution of every possible initial state to properly understand a dynamical system. In practice, if the equations of motion are known, then one resorts to numerical simulation and sampling trajectories from several initial states. Experimentally, however, it is often the case that one is in possession of only a single long orbit. In addition, there is often no direct access to the state variables themselves as the time series may comprise values of a measurement function applied to the state of the system.

A. Symbolic itinerary

Let $\{x_n^*\}_{n=1}^N$ denote an arbitrary scalar time series of N temporally ordered measurements of the state of a dynamical system. We require that these values are drawn from a linearly ordered set. Our main assumption is that the *observable*, i.e., the sequence x_n^*, represents the trajectory of a deterministic system. It may correspond to one of the state variables $\{x_j\}_{j \in \{1,\ldots,d\}}$ or to a measurement function of the full state of the system $g(\boldsymbol{x})$. For simplicity we drop the asterisk hereinafter.

The ordinal symbolization technique [22] consists of a coarse-grained segmentation of the dataset into elements (windows) of equal length m, i.e., akin to embedding the time series to m-dimensional space using a time delay τ. Successive windows are *not* overlapping at $w = w(m) \in \{1, m\}$ points. Thus, the maximum number of windows occurs in the case of maximal overlap at $m - 1$ points ($w = 1$) and is equal to $N - m + 1$. However, the minimum number occurs in the nonoverlapping case ($w = m$) where the total number of windows equals $\lfloor \frac{N}{m} \rfloor$. To illustrate the difference between parameters τ and w, consider the first two segments obtained by setting $m = 3$, $\tau = 2$ and (A) $w = 1$ or (B) $w = 4$. Time

lag τ dictates the relation between points *within* a segment, hence the first two segments are given by (x_1, x_3, x_5) and (x_2, x_4, x_6) in scenario A. Slide lag w dictates the relation between successive *segments*, in particular the extent to which they do not overlap. Hence, the first two segments in scenario B are given by (x_1, x_3, x_5) and (x_5, x_7, x_9). The first component of consecutive segments under scheme A is generated by the sequence $x_1, x_2, x_3, \ldots$ due to $w = 1$, whereas the corresponding sequence under scheme B is $x_1, x_5, x_9, \ldots$ due to $w = 4$.

Denote the resulting collection of segments (or embedding vectors) by an indexed sequence of m-tuples of the form

$$\boldsymbol{x}_i^{(m)} = (x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}) \in \mathbb{R}^m$$

with indices $\quad i = 1, 1+w, 1+2w, \ldots, \left\lfloor \dfrac{N-m}{w} \right\rfloor + 1,$

preserving the temporal order of windows. Each window is mapped to a symbol drawn from a finite alphabet. The set of permutations of the natural numbers $\{1, 2, \ldots, m\}$, over which the symmetric group $\mathcal{S}_m$ acts, will serve as the alphabet of choice to create a *symbolic ordering* for each window which reflects the ordinal relations between points within. Ordinal symbols can be thought of either as the set of re-orderings of $\{1, 2, \ldots, m\}$ or as the group elements of $\mathcal{S}_m$ acting on the identity element through the symmetric group's operation, i.e., composition of permutations; the original choice of the identity element as a reference frame is arbitrary and any member of $\mathcal{S}_m$ comprises a valid alternative.

The mapping may be defined in two equivalent ways, namely *chronological index ranking* introduced by Bandt and Pompe [22] and *amplitude ranking* introduced by Small [7]. According to the former, shifted time indices within a window are stored in a vector indexed by the corresponding amplitude rank in ascending order. For example, consider one of the windows obtained from segmentation into windows of length $m = 4$ with $\tau = 1$, $w = 1$, say the seventh $\boldsymbol{x}_7^{(4)} = (x_7, x_8, x_9, x_{10})$, whereby relative magnitude relations are given by

$$x_{10} < x_8 < x_7 < x_9.$$

By shifting indices $\{7, 8, 9, 10\}$ to the set $\{1, 2, 3, 4\}$, the associated permutation is (4,2,1,3), or in cyclic notation (143)(2). This is due to the fact that the first point of this segment chronologically (x_7) is the second largest value in the group and so assumes the *third* position (due to order being ascending). Similarly, the second chronological value (x_8) is the third largest and so placed in the *second* position, x_9 (third) is the largest and so assumes the *fourth* position, etc. Thus, according to this mapping, time indices of sample points are sorted.

However, according to amplitude ranking, points within a window are ranked in terms of relative amplitude in descending order, i.e., time indices remain fixed and each sample point is simply assigned a rank. For instance, the 4-tuple (x_7, x_8, x_9, x_{10}) is assigned the permutation (2,3,1,4), or (123)(4), since x_7 is the *second* largest member of the group, x_8 is the *third* largest, x_9 is the largest, etc. The sketch in the top panel of Fig. 2 illustrates the difference between the two for the case $m = 3$. It can be easily shown that both formulations are equivalent by defining a bijective map
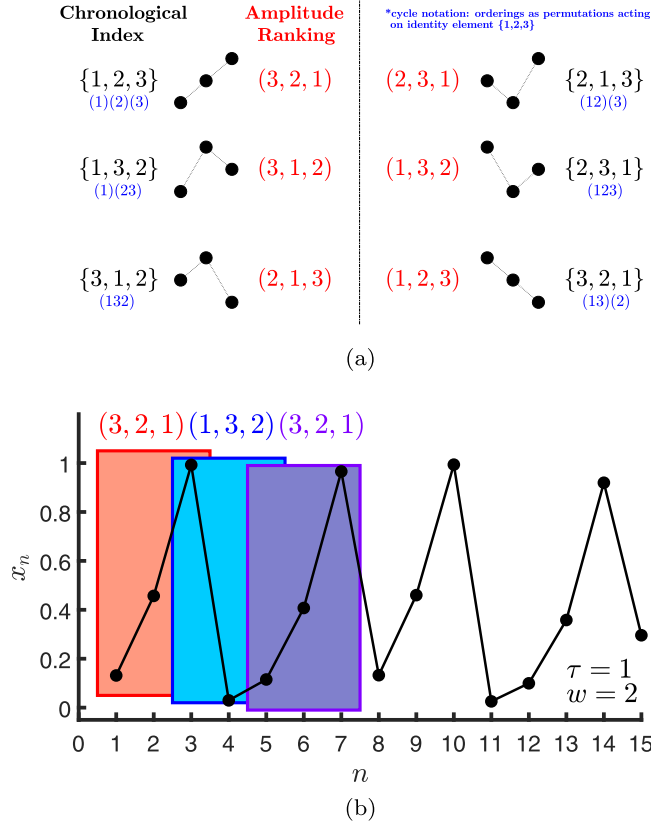
(a)

(b)

FIG. 2. (a) Juxtaposition between chronological index rank and amplitude ranking. (b) Assignment of ordinal patterns of length $m = 3$ to a chaotic time series generated by the logistic map $x_{n+1} = 4x_n(1 - x_n)$ with $x_n \in [0, 1]$ via amplitude ranking.

between the two. In the case $m = 2$, this rank is equivalent to the traditional setting in symbolic dynamics, i.e., mapping of trajectories to infinite sequences of elements in the binary set $\{0, 1\}$.

Formally, an *ordinal pattern* of length m for the window labeled by $\boldsymbol{x}_i^{(m)}$

$$\boldsymbol{\pi}(\boldsymbol{x}_i^{(m)}) = (\pi_1, \pi_2, ..., \pi_j, ..., \pi_m), \quad \pi_j \in \{1, ..., m\}$$

may be defined as the permutation $\boldsymbol{\pi} \in \mathcal{S}_m$, which arranges the points within according to their order, i.e., $x_{\pi_1} < x_{\pi_2} < \cdots < x_{\pi_m}$ (chronological index rank), or as the transformation $\boldsymbol{\pi} : \mathbb{R}^m \to \mathcal{S}_m$ such that x_{i+j-1} is the π_jth largest point in $\boldsymbol{x}_i^{(m)}$ (amplitude rank). In the event that two elements of $\boldsymbol{x}_i^{(m)}$ are equal, we arbitrarily pick the one that occurs first chronologically as the smallest [22].

By mapping every window to an ordinal pattern, one obtains a symbol sequence corresponding to a given time series *uniquely*. We represent the *ordinal symbolic itinerary* of time series $\{x_n\}_{n=1}^N$ by

$$\{s_n\}_{n=1}^{\lfloor \frac{N-m}{w} \rfloor + 1}, \quad s_n \in \mathcal{S}_m \ \forall n \in \mathbb{N}. \tag{4}$$

The empirical probability distribution of ordinal patterns is defined via their relative frequency of occurrence within s_n,

$$\text{Prob}[\boldsymbol{\pi}; s_n] = \mathbb{P}[\boldsymbol{\pi}] = \frac{|\{n | s_n = \boldsymbol{\pi}\}|}{\lfloor \frac{N-m}{w} \rfloor + 1}, \quad \boldsymbol{\pi} \in \mathcal{S}_m, \tag{5}$$
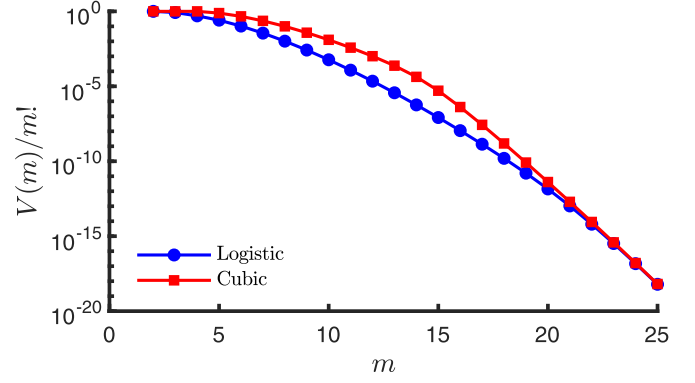

FIG. 3. Proportion of admissible ordinal patterns as m increases. Estimates computed from scalar time series generated via iteration of the expansive chaotic maps given by the (a) logistic equation $x_{n+1} = 4x_n(1 - x_n)$ with $x_n \in [0, 1]$ and (b) cubic equation $x_{n+1} = 3x_n - 4x_n^3$ with $x_n \in [-1, 1]$.

where $|\cdot|$ denotes cardinality of a set and "|" means "such that."

The number of admissible (*distinct* occurring) patterns is

$$\mathcal{N}(s_n, m) = |\{\boldsymbol{\pi} \in s_n | \mathbb{P}[\boldsymbol{\pi}] \neq 0\}|. \tag{6}$$

Since the cardinality of $\mathcal{S}_m$ is $m!$, one can evaluate the relative count of admissible patterns computed from a time series over the set of all possible, $\mathcal{P}_a^{(m)}(x_n)$, and the corresponding proportion in the complementary set of forbidden ordinal patterns (patterns which *cannot* be realised in a time series due to deterministic constraints [56]). The relative count of forbidden patterns equals

$$\mathcal{P}_f^{(m)}(x_n) = 1 - \mathcal{P}_a^{(m)}(x_n) = 1 - \frac{1}{m!} \mathcal{N}(s_n, m). \tag{7}$$

In practice, $\mathcal{P}_a^{(m)}, \mathcal{P}_f^{(m)}$ are both estimators of the true numbers in the case of $N < \infty$, i.e., when we are in possession of a trajectory but not the full orbit. Note that the absence of an m-pattern pervades all longer patterns for larger m in the form of *outgrowth forbidden patterns* whose growth is super-exponential in contrast to admissible m-patterns (exponential growth with m). The semilogarithmic plot depicted in Fig. 3 illustrates the rate of decrease of the proportion of admissible patterns computed from chaotic time series ($N = 10^6$) of the logistic (black circles) and cubic (red squares) maps. Note that the rate is faster than exponential. This has been theoretically described for discrete maps [56]. In addition, the discrepancy in the rate corresponding to each map is due to the inherently different dynamics. Since both maps are strongly mixing, the larger state space $[-1, 1]$ is expected to produce a greater diversity of ordinal patterns for finite m. For $m \geqslant 23$, the two curves coincide due to undersampling effects.

B. Ordinal partition

In the previous section, we outlined the details of the generalized BP methodology for ordinal symbolization of a given time series. Here we describe the fundamental construction for the modeling process we propose in this study. To speak of Markovian descriptions of deterministic dynamics, we shall restrict our attention to even more specific families of systems.

The characteristic shared by all systems under study is *chaotic* behavior, i.e., (I) sensitivity to initial conditions, (II) topological mixing, and (III) the existence of dense periodic orbits (Ch. 1, Sec. 1.8, Def. 8.5 [58]). In addition, the evolution rule ϕ is required to be a measurable transformation. Endow $\mathcal{M}$ with a σ-algebra $\Sigma_{\mathcal{M}}$ and associate with an invariant probability measure

$$\nu : \mathcal{M} \to [0, 1] | \nu(A) = \nu[\phi^{-1}(A)] \quad \forall A \in \Sigma_{\mathcal{M}}. \quad (8)$$

The probability space denoted by the triplet $(\mathcal{M}, \Sigma_{\mathcal{M}}, \nu)$ is also equipped with a *filtration* $\{\mathcal{F}_{n \geq 0}\}$, i.e., an increasing sequence of σ-algebras on the measurable space $\mathcal{M}$ such that $n_1 \leq n_2 \Rightarrow \mathcal{F}_{n_1} \subseteq \mathcal{F}_{n_2}$. Since ν is ϕ-invariant, we speak of a *measure-preserving* dynamical system, i.e., one which obeys a 'conservation of mass' principle with respect to the action of the evolution operator ϕ. Infinitely many invariant measures are typically admitted by deterministic chaotic dynamics. Experimentalists often wish that ν corresponds to the *natural* invariant measure, i.e., the distribution of state space mass over a typical orbit of the system. Therefore, our focus is restricted to examination of ergodic systems [59]. The measure ν satisfies

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\phi^n(\boldsymbol{x})) = \int_{\mathcal{M}} f \, d\nu \quad (9)$$

for any continuous function $f : \mathcal{M} \to \mathbb{R}$ and Lebesgue-almost all initial conditions $\boldsymbol{x} \in \mathcal{M}$.

Given that the natural measure of the system is unknown, one resorts to simpler alternatives to compute an estimate thereof from x_n. This inadvertently leads to the consideration of different types of finite state-space partitions (Sec. 6.1 [30]). In practice, the best alternative for the estimator of ν, which we denote by $\hat{\nu}$, reduces to essentially counting symbols using one of several alternatives for a state-space partition [30,32,44]. This finite set of symbols represents a partition of state space into a finite number of connected regions of nonempty interior. Usually the elements of the partition take the form of regular-grid cells [32] whose collection covers $\mathcal{M}$. Dynamics occurring within each element are completely ignored, while the model captures coarse-grained dynamics between elements. It is this classical framework that we adopt, but by adapting it to our setting. In place of $\hat{\nu}$ we choose the normalized Lebesgue measure.

Since ordinal symbols constitute our choice, the first step of our proposed procedure amounts to a mapping of $\mathcal{M}$ to a higher-dimensional space $\mathbb{R}^m$ with $m > d$ via a time-delay embedding of a single point lag [60]. Provided certain smoothness and genericity conditions on ϕ which are fulfilled by the family of systems defined above, a sufficiently large m—specifically larger than twice the fractal (box-counting) dimension D_0 of manifold $\mathcal{M}$ [61]—ensures the existence of a diffeomorphism between the original and reconstructed state spaces. The image state space $\mathcal{M}^{\mathrm{IMG}}$, and not the original manifold $\mathcal{M}$, is the space which will be partitioned into a finite number of elements. Therefore, topological equivalence between original and embedded dynamics is ensured via counting patterns of length larger than twice the fractal dimension of the underlying dynamics.
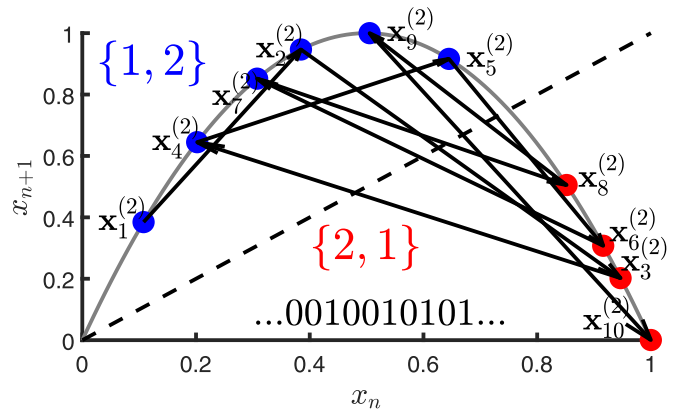

FIG. 4. Chaotic logistic trajectory of Fig. 3 embedded in $\mathbb{R}^2$ and corresponding ordinal partition of image state space $\mathcal{M}^{\mathrm{IMG}} = [0, 1] \times [0, 1]$ into elements Q_1, Q_2 indexed by (1,2) and (2,1), respectively. Arrows represent temporal order. Representing Q_1, Q_2 by 0 and 1, respectively, leads to the binary itinerary corresponding to this trajectory.

The image state space $\mathcal{M}^{\mathrm{IMG}} \subseteq \mathbb{R}^m$ may then be covered by a finite collection of connected sets of nonempty interior, $\mathcal{Q} = \{Q_1, Q_2, \ldots, Q_{m!}\}$, indexed by members of $\mathcal{S}_m$, and referred to as an *ordinal partition* of *order m*. Recast here, we have that given any point $\boldsymbol{x} = (x_1, \ldots, x_m)^{\mathrm{T}} \in \mathcal{M}^{\mathrm{IMG}}$,

$$\boldsymbol{x} \in Q_{\boldsymbol{\pi} \in \mathcal{S}_m} \Leftrightarrow x_{\pi_1} \leq x_{\pi_2} \leq \cdots \leq x_{\pi_m}, \quad (10)$$

which must always be true for some permutation label $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_m) \in \mathcal{S}_m$ as we are only concerned with linearly ordered sets. Moreover, this mapping is obviously *unique*. The lack of strict inequalities connotes the resolution of ranking ties via chronological order.

Intuitively, each element of an ordinal partition is spatially defined by a region whose boundaries are formed by the $x_i = x_j$ cross-sectional hyperplanes of Euclidean space $\mathbb{R}^m$ (x_i denotes the ith coordinate). For instance, in $\mathbb{R}^2$ the sole boundary is provided by the identity line $x = y$ as Fig. 4 demonstrates for the chaotic logistic trajectory. In $\mathbb{R}^3$ these are the $x = y$, $x = z$, and $y = z$ planes as depicted in the partition of Fig. 5. Notice that the identity line $x = y = z$ is a shared boundary between all $3! = 6$ regions (different color in Fig. 5). In $\mathbb{R}^4$ the six three-dimensional spaces $x_i = x_j$ where $i, j = 1, \ldots, 4$ segment the full space into $4! = 24$ regions with the identity line $(1, 1, 1, 1)^{\mathrm{T}}$ again being a common boundary, a fact which holds in general regardless of the value of m. Therefore, each of the $m!$ regions in $\mathbb{R}^m$ may also be prescribed by a hyper-arc denoting a different angular location of points with respect to the identity line $\boldsymbol{1} = (1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^m$. The pattern length dictates both the resolution of the partition (uniform size $1/m!$) and its dimensionality. Upon increasing the resolution, data is mapped to increasingly higher-dimensional spaces irrespective of the original dimensionality.

This is not a partition in the strict sense as we are not faced with pair-wise disjoint elements due to shared boundaries. To overcome this difficulty in practice, Bandt and Pompe [22] suggested using chronological order as a criterion for sorting ranking ties. This is equivalent to sending every observed point on the boundary into the interior of one of the two
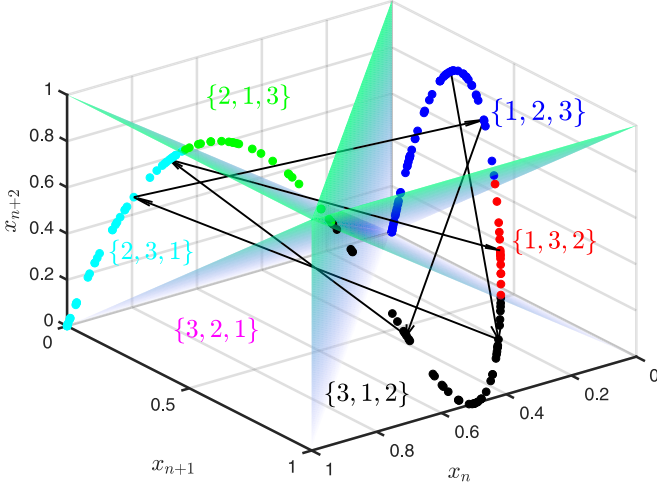
FIG. 5. Chaotic logistic trajectory of Fig. 3 embedded in $\mathbb{R}^3$ and corresponding ordinal partition ($m = 3$) of image state space $\mathcal{M}^{\text{IMG}}$ into elements $Q_1, ..., Q_6$. Arrows represent temporal order. The region indexed by the permutation $(3,2,1)$ is never visited by the trajectory; this pattern is forbidden.

adjacent elements. However, in rigorous terms, the covering has the following two properties:

(I) $\cup_{i=1}^{m!} Q_i = \mathcal{M}^{\text{IMG}}$,

(II) $\text{Int}(Q_{\boldsymbol{\pi}}) \cap \text{Int}(Q_{\boldsymbol{\xi}}) = \varnothing$ if $\boldsymbol{\pi} \neq \boldsymbol{\xi}$,

where Int denotes the interior of a set, excluding boundary points. The set of elements of the partition $\mathcal{Q}$, labeled by permutations of natural numbers, will serve as the symbolic alphabet of an associated Markov model for the evolution rule $\boldsymbol{\phi}$ once our symbolic dynamics are defined as follows.

Given a finite time series such as x_n, we possess a temporally ordered trajectory that is a subset of the true dynamical orbit $\mathcal{O}(\boldsymbol{x_0})$. Embedding with a sufficiently high m and $\tau = 1$ allows us to obtain a subset of a topologically equivalent orbit $\mathcal{M}^{\text{IMG}}$, designated by the sequence

$$\mathcal{O}(\boldsymbol{x_0^{(m)}}) = \{\ldots, \phi_*^{-3}(\boldsymbol{x}^{(m)}), \phi_*^{-2}(\boldsymbol{x}^{(m)}), \phi_*^{-1}(\boldsymbol{x}^{(m)}), \boldsymbol{x}^{(m)},$$
$$\phi_*(\boldsymbol{x}^{(m)}), \phi_*^2(\boldsymbol{x}^{(m)}), \phi_*^3(\boldsymbol{x}^{(m)}), \ldots\},$$
$$\forall \boldsymbol{x}^{(m)} \in \mathcal{O}(\boldsymbol{x_0^{(m)}}),$$

whereby ϕ_* topologically conjugate to ϕ. Since $\mathcal{M}^{\text{IMG}}$ is fully covered by partition $\mathcal{Q}$, which is guaranteed by property (I), every point $\boldsymbol{x}^{(m)} \in Q_{\boldsymbol{\pi}}$ for some permutation $\boldsymbol{\pi}$. Similarly for all its iterates $\phi_*^n(\boldsymbol{x}^{(m)})$. Consider a sequence comprising of labels of $Q_n \in \mathcal{Q}$ which is defined via the orbit $\mathcal{O}(\boldsymbol{x_0^{(m)}})$, i.e., a bi-infinite sequence

$$\{s_n\}_{-\infty}^{\infty} = \{s_n \in \mathcal{S}_m | \phi_*^n(\boldsymbol{x}^{(m)}) \in Q_{s_n}\}, \qquad (11)$$

which represents the analogous *symbolic orbit*. The finite symbolic itinerary (abuse of notation since we only refer to the finite itinerary s_n hereinafter) s_n defined in Eq. (4) constitutes a time-contiguous sample drawn from this sequence. Using a decimal point, we can separate $s_{i<0}$ from $s_{i\geq 0}$ in the form $\ldots s_{-2}s_{-1}.s_0 s_1 s_2 s_3 s_4 \ldots$ and speak of current state, future and past. This construct suffices for delineation of shifts (and subshifts) of finite type and the associated symbolic dynamics.

IV. MARKOVIAN FRAMEWORK

Instead of following the direction of topological conjugacy and studying subshifts of finite type, as in Ref. [62] within the Amigó paradigm, we opt for a Markovian description. We define discrete-time stochastic processes of countable state space for increasing m on the set of permutations by means of the counting measure and information extracted from a sample trajectory. The procedure is defined via the counting measure applied to the symbolic itinerary s_n and the associated σ-algebra $\Sigma_{\mathcal{Q}}$ of all subsets of the symbol space. In particular, we record pairs of consecutive ordinal patterns drawn from the ordered product space $\mathcal{S}_m \times \mathcal{S}_m$. They represent *forward-time transitions* between admissible patterns, or alternatively a single action of ϕ_* on elements of the partition $\mathcal{Q}$. Given a consecutive pair of ordinal patterns in the symbolic itinerary, say $\ldots, s_n = \boldsymbol{\pi}, s_{n+1} = \boldsymbol{\xi}, \ldots$ denoting the corresponding partition elements $Q_{\boldsymbol{\pi}}, Q_{\boldsymbol{\xi}}$, we can compute $\mu(Q_{\boldsymbol{\pi}} \cap \phi_*^{-1} Q_{\boldsymbol{\xi}})$—the number of times the itinerary is in state $Q_{\boldsymbol{\pi}}$ and the ϕ_*-preimage of $Q_{\boldsymbol{\xi}}$ simultaneously. This quantity represents the proportion of mass moving from region $Q_{\boldsymbol{\pi}}$ to region $Q_{\boldsymbol{\xi}}$ in state space under a single action of the dynamics. Division by $\mu(Q_{\boldsymbol{\pi}})$, the normalized version of which is the probability measure on the space of ordinal patterns as defined in Eq. (5), leads to the required Markovian transition probabilities. Given probability space $(\mathcal{Q}, \Sigma_{\mathcal{Q}}, \mu)$ along with filtration $\{\mathcal{F}_n\}$, we define the $\mathcal{S}_m$-valued stochastic process $S = (S_n, n \in \mathbb{N})$ which possesses the Markov property, i.e., $\forall \boldsymbol{\pi} \in \mathcal{S}_m$ and n_1, n_2 with $n_1 < n_2$,

$$\mathbb{P}\big[S_{n_2} = \boldsymbol{\pi} | \mathcal{F}_{n_1}\big] = \mathbb{P}\big[S_{n_2} = \boldsymbol{\pi} | S_{n_1}\big], \qquad (12)$$

if the input time series comprises a measurement function of a signal generated by a deterministic source. The most fundamental question is that of convergence as $m \to \infty$ since the number of admissible patterns grows exponentially with m [56]. However, the Shannon-McMillan-Breiman theorem asserts that these entropies do not diverge as m becomes large provided the source is a discrete-time ergodic stationary process [63]. That is, the ordinal Markov chain defined in Eq. (12) is characterized by the *asymptotic equipartition property*.

Note that, strictly speaking, S obeys the Markov property given by Eq. (12) only if $w \geq m$; in the case of inequality, some information loss occurs since there are time series points that are not encoded. This constitutes a *nonoverlapping scheme*. If $w < m$, then successive patterns overlap at $m - w$ points, with $w = 1$ producing the *maximally overlapping* scheme. Hence, successor patterns are conditioned on their predecessor, i.e., each new pattern contains part of the previous pattern and becomes part of the following pattern. Therefore, the constraint imposed on the set of admissible ordinal transitions leads to Markov chain S having a finite memory. This setup can be treated as a Markov chain of order $m - w$, with $w \leq m - 1$ finite. The future state depends on the past $m - w$ states.

A. Ordinal networks

Ordinal symbolization so as to obtain itinerary s_n together with Markov chain S allow the association of a given time

series with a directed graph, as is common in the study of Markov chains of countable state (Ch. 1 [64], Sec. 1.4, 1.7 [65]). This is what we term an *ordinal network*. It is constructed as follows. The set of observed admissible patterns is mapped onto the set of nodes $\mathcal{V} = \{1, 2, 3, \ldots, V\}$ of a graph, say $G(\mathcal{V}, \mathcal{E})$. The total number of nodes (assumed to be nonisolated by construction) equals the *network size* $|\mathcal{V}| = V(m)$, is equal to $\mathcal{N}(s_n, m)$. Directed edges are assigned to pairs of patterns occurring in temporal succession, i.e., the set of links (or edges) $\mathcal{E}$ corresponds one-to-one to admissible forward-time transitions between regions of the ordinal partition $\mathcal{Q}$ as observed in the given time series.

Formally, bijective mappings $\kappa : \mathcal{S}_m \to \mathcal{V}$ and $\psi : \mathcal{S}_m \times \mathcal{S}_m \to \mathcal{E}$ are applied to the set of admissible symbols (and the product set of ordered pairs of symbols) of itinerary s_n to procure an OPN topology. Nonelementary events in the context of the previously defined ordinal stochastic process of order m, say of the form $Q_\pi \cap \phi_*^{-1} Q_\xi$, are represented by a link labeled ij, whereby it is understood that $\exists n \in \mathbb{N}$ such that

$$[\kappa(s_{n-1}), \kappa(s_n)] = (\pi, \xi) \bigcap \psi(s_{n-1}, s_n) = (\pi, \xi) = (i, j),$$
(13)

where $\pi, \xi \in \mathcal{S}_m$ are two ordinal symbols which appear successively within the symbolic itinerary. They, respectively, form the labels of partition elements Q_π, Q_ξ and nodes indexed by $i, j \in \mathcal{V}$. Thus, we can revert to the familiar i, j notation following this, which facilitates comparison of our formulas to the existing literature on ergodic theory.

The elements of the adjacency matrix $\boldsymbol{W} = (W_{ij})_{i,j=1}^{\mathcal{N}(s_n, m)}$ form a *directed multigraph* and are simply given by the frequencies

$$W_{ij} = \mu(Q_i \cap \phi_*^{-1} Q_j),$$
(14)

where μ denotes the Lebesgue measure given infinitely long time series. In practice, it corresponds to the counting measure. Directed multigraphs comprise a natural representation for a random walk realised according to the rules specified by finite-state time-homogeneous Markov chains such as S.

The associated *transition graph* dictates that edges are normalized by the number of nearest (outgoing) neighbors of each node; i.e., its elements are identical to the transition probabilities of chain S. Then a right-stochastic matrix, $\boldsymbol{T} = (T_{ij})$, is produced with rows summing up to unity. This corresponds to the weighted state-transition matrix of the Markov chain. It is defined by the conditional probability

$$T_{ij} = \frac{\mu(Q_i \cap \phi_*^{-1} Q_j)}{\mu(Q_i)}.$$
(15)

We interpret entry T_{ij} as the probability that a typical point in region Q_i of embedding space moves into Q_j under one iteration of the evolution operator ϕ_*.

The binary formulation $\boldsymbol{A} = (A_{ij})$ given by

$$A_{ij} = \mathbb{1}_{(0,\infty)}[\mu(Q_i \cap \phi_*^{-1} Q_j)],$$
(16)

where $\mathbb{1}_{(0,\infty)}(x)$ denotes the indicator function on an open semi-infinite interval, is the corresponding *connectivity matrix*. It constitutes an unweighted directed graph with $A_{ij} = 1$ if a forward-time transition between the corresponding ordinal symbols is observed in s_n, otherwise it is equal to 0.
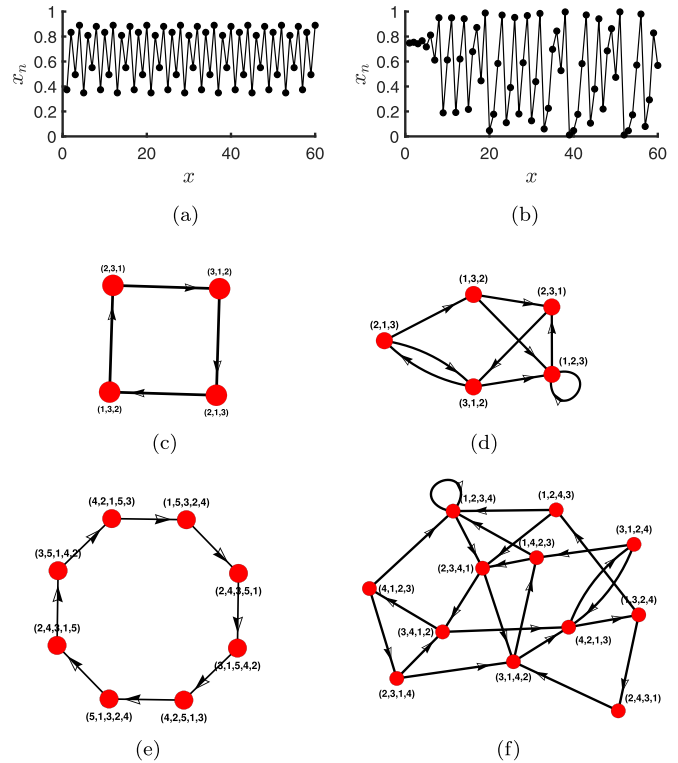

FIG. 6. (a), (b) Example trajectories from an 8-periodic ($r = 3.56$; left) and a chaotic ($r = 4$; right) dynamical regime of the logistic map $x_{n+1} = r x_n (1 - x_n)$. (c), (d) Respective ordinal networks using patterns of length $m = 3$, single-point time delay ($\tau = 1$), and maximal overlap ($w = 1$). (e), (f) OPNs for higher values of m.

We shall show that matrix $\boldsymbol{A}$ captures important topological information about the underlying dynamics behind the input time series.

To summarize, constructing OPNs from a specified time series comprises a multistage process. First, a nonlinear transformation to a different set of coordinates in a higher-dimensional space is applied. Second, this new state space is partitioned into a finite collection of disjoint sets by means of location of points relative to the identity line $x = y = z = \ldots$. Third, a subsequent mapping to a symbol space is performed. Fourth, this allows the formulation of a discrete-time finite-valued Markov chain. Directed and weighted networks correspond to the natural transition graph representation thereof (see Fig. 6 for example networks representing periodic and chaotic dynamics). All relevant information is extracted via application of the counting measure to the symbolic itinerary s_n.

B. Ergodicity

Ergodic theory is concerned with the statistical properties of deterministic dynamical systems which admit an invariant measure. In abstraction, the field concerns qualitative properties of the action of a group on a measure space. This constitutes a departure from the Lagrangian perspective of tracking the dynamics of an individual point and an adaptation of a viewpoint of the motion of collections of particles (or *densities*). Intuitively, one considers the evolution of a "mass

of particles" in state space; the aim being the study of the temporal change in the 'size' of a specified set of particles. The appropriate mathematical framework for general $\mathcal{M} \subseteq \mathbb{R}^d$ is via the Borel σ-algebra (smallest subset of the set of all measurable subsets of $\mathbb{R}^d$), whereby the notion of "size" may be quantified in a meaningful sense by means of the Lebesgue measure.

Dynamical systems may be classified depending on whether (i) *mass* or (ii) *volume* is preserved in state space under the dynamics. In the former case, the ergodicity property asserts that, under certain conditions, the time average of a function along a single trajectory exists almost everywhere and is related to the spatial average. This property, which formalises the notion of *typicality*, is guaranteed by strong analytical results such as the ergodic theorems by Birkhoff [59] and Khinchin [66]. In the latter class, the distinction is between *conservative* (area/volume/Lebesgue measure-preserving) dynamical systems, *dissipative* (volume-contracting) systems and *expanding maps* (volume-generating). Note that the latter two are still measure preserving, but not of the Lebesgue measure. All three such types of systems are amenable to a Markovian approach.

A major advantage of following the Markov modeling paradigm is the ability to overcome all the drawbacks associated with simulating (or examining) a long orbit of a dynamical system, which comprises the traditional approach [30]. Namely, that orbits may display nonequilibrium behavior for a long time before settling into an asymptotic mode and computer round-off errors or limited experimental precision. Rather than compounding a sequence of iterations, one can treat a time series generated by an ergodic system as a collection of *single iterates* of the governing dynamical law via a stochastic approximation provided by the Markov model.

The framework we propose for interpreting OPNs is essentially an adaptation of the work by Froyland [31] and Dellnitz and Junge [32] on the discretization of the Perron-Frobenius operator via Ulam's method, but in embedding space rather than the original state space. In one-dimensional systems, Li [67] and Boyarsky [68] first demonstrated how powerful this approach can prove. Froyland [69] extended the former's analytical results to random interval maps. The main idea lies on the fact that the statistical properties of interest can be rendered stable by the addition of a small amount of noise [30]. Thus, instead of analyzing the deterministic dynamics directly, one may wish to slightly perturb and create a Markov model. The details of the method for approximating a natural measure in this fashion have been outlined in Ref. [31].

C. Strongly connected graph and irreducibility

The pertinent questions for any Markov chain are the existence and uniqueness of a stationary distribution, which reflects the long-term average occupation of each state if the underlying process is ergodic. Existence is generally guaranteed by finiticity and time-homogeneity (Sec. 1.5, Ch. 4, Convergence Theorem 4.9 [65]) of the chain in question, while uniqueness is slightly trickier. The property of irreducibility

is a facet of OPNs which can be instantly checked after symbolization. Topological transitivity is expected to lead to irreducible Markov chains, however finiticity can alter the produced network in the case of undersampling. In the case of lack of irreducibility due to finiticity, we propose a quick and benign way of modifying the network structure which ensures this property. First, we proceed with the necessary definitions.

State (node) j is said to be *accessible* from state i if $\mathbb{P}(S_{n_{ij}} = j | S_0 = i) > 0$ for some $n_{ij} \in \mathbb{N} \cup \{0\}$, i.e., there is a nonzero probability that a system realisation (or alternatively, a random walk on the transition graph $\boldsymbol{T}$) starting from node i will reach node j at some point in the future (dependent on both i and j, hence the index). Transition ij is represented by a self-loop in $\boldsymbol{T}$ if $n_{ij} = 0$. A Markov chain is *irreducible* if all states are accessible from all others. Formally,

$$\exists n \in \mathbb{N} | \mathbb{P}(S_n = j | S_0 = i) = \boldsymbol{T}_{ij}^n > 0 \quad \forall i, j \in \mathcal{V}. \quad (17)$$

Furthermore, a state i has period $k \geqslant 1$ if recurrent trajectories *must* occur at integer multiples of k, i.e., $k = \text{GCD}\{n \in \mathbb{N} | \mathbb{P}(S_n = i | S_0 = i) > 0\}$. The acronym GCD stands for greatest common divisor. The chain is *aperiodic* if $k = 1 \; \forall i \in \mathcal{V}$. These two notions are essential for the facts we establish as follows.

First of all, if a finite-state Markov chain is irreducible and aperiodic, there exists a *unique* stationary distribution $\boldsymbol{p} = (p_1, p_2, \ldots, p_V)$ which satisfies the axioms of probability (nonnegativity, $\sum_i p_i = 1$ and countable additivity) and such that

$$\boldsymbol{p}\boldsymbol{T} = \boldsymbol{p}. \quad (18)$$

This is the positive invariant measure of the Markov chain. This statement and convergence to equilibrium are well-known facts, see, e.g., Thm. 1.7.2, Thm. 1.7.7, Thm. 1.8.3 in Norris [64]. Eq. (18) is identical to the left-eigenvector equation of $\boldsymbol{T}$ with eigenvalue equal to unity. The appropriate multiple in this eigendirection corresponding to $\boldsymbol{p}$ is obtained by normalizing the *leading* left eigenvector. (Since the Perron-Frobenius theorem dictates that the unit eigenvalue is the largest for a stochastic matrix, it is simple, and so the corresponding eigenspace is one-dimensional and all elements of the leading eigenvector are positive.)

Second, we are also concerned with the property of *ergodicity*. The celebrated ergodic theorem due to Birkhoff [59] was initially proved within the framework of dynamical systems on smooth manifolds, but was later recast in terms of a general information source by Khinchin [66]. It is this form we are interested in since we would like to ensure that the ordinal Markov chain S has a representative limiting behavior in terms of a time-average from a long trajectory. Fortunately, irreducibility suffices to prove this fact (Thm. 1.10.2 [64]).

Third, irreducibility is manifested in the form of a *strongly connected* transition graph, i.e., a path exists from every node to any other. We remark that an OPN will be irreducible by virtue of construction since a node is connected to the nodes corresponding to the predecessor and successor symbols within the itinerary s_n, except in the following two problematic cases: (a) the node corresponding to s_1 has no incoming links, or (b) the node corresponding to $s_{\lfloor \frac{N-m}{w} \rfloor + 1}$ has no outgoing links. This event becomes more likely with

increasing m, however it is quite rare judging from numerous numerical simulations and never occurs unless the symbolic itinerary or the OPN is severely undersampled for specified m. In addition, we propose the simple fix of removing the first or last value of a time series, repeating the irreducibility check (via strong component detection algorithms) and iterating this procedure if necessary.

Finally, we remark that it can be easily shown that the relative (over all nodes) out-strength distribution of the directed multigraph W and the stationary distribution of the ordinal Markov chain S are identical. Denote by $k_W^{out}(i)$ the *out-strength* of node $i \in \mathcal{V}$, i.e., the total number of adjacent outgoing links

$$k_W^{out}(i) = \sum_{j=1}^{V} W_{ij}. \tag{19}$$

This term is equivalent to the *weighted out-degree* of a node. By "out-degree" we refer to the (binary) connectivity graph represented by A. It corresponds to the number of *distinct* outgoing neighbors, i.e., the total number of successor patterns.

Note that by nature of construction of an ordinal multigraph, every node has the same number of incoming and outgoing links with the exception of possibly two nodes. This is a direct consequence of the fact that all the symbols in itinerary s_n, except the first and last, are connected to their predecessor by an incoming transition and their successor by an outgoing transition. Otherwise we have that $k_W^{in}(i|s_1 = i) = k_W^{out}(i|s_1 = i) - 1$ and $k_W^{in}(i|s_{\lfloor \frac{N-m}{w} \rfloor + 1} = i) = k_W^{out}(i|s_{\lfloor \frac{N-m}{w} \rfloor + 1} = i) + 1$. Provided a sufficiently long time series, equality follows in the limit $k_W^{out}(i) \gg 1$. Observe that the out-strength distribution of the network obeys Eq. (18), the left-eigenvector equation which element-wise takes the form

$$\sum_i p_i \cdot T_{ij} = p_j \quad \forall i, j \in \mathcal{V}. \tag{20}$$

Consider the left-hand side of Eq. (20)

$$\sum_i k_W^{out}(i) T_{ij} = \sum_{i|A_{ij}=1} k_W^{out}(i) \frac{W_{ij}}{k_W^{out}(i)}$$

$$= \sum_{i|A_{ij}=1} W_{ij} = k_W^{in}(j) = k_W^{out}(j). \tag{21}$$

Evidently, the out-strength (and in-strength) of multigraph W is a left-eigenvector and can be made identical to the stationary distribution p via normalization. Figure 7 shows the Ikeda attractor ($\mu = 0.9$) with state-space points of the chaotic trajectory colored according to the relative out-strength of the corresponding node in the OPN computed for $m = 10$. Since the nodal out-strength corresponds to a symbol s_i in the itinerary, we mapped colors to the first value x_i in the window $x_i^{(m)}$.

V. NUMERICAL RESULTS

A. Periodicity in ordinal symbolic itineraries

The first test that the OPN construction is subjected to is related to its capacity for accurate representation of periodicity. We examine network topology separately for the case of (a) discrete-time and (b) continuous-time systems, as sampling frequency plays a central role in the latter case.
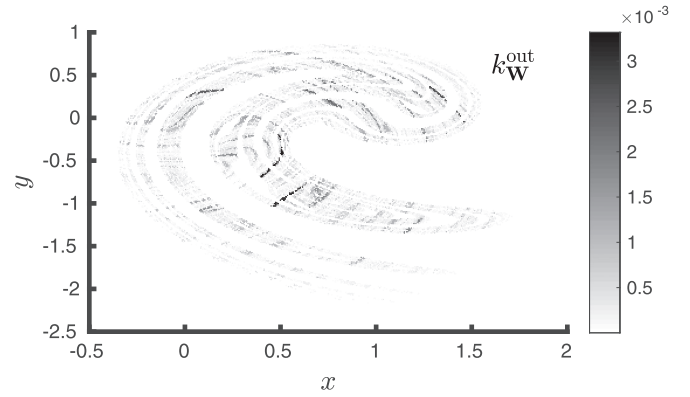

FIG. 7. Chaotic attractor generated by setting $\mu = 0.9$ in the Ikeda map $x_{n+1} = 1 + \mu(x_n \cos \theta_n - y_n \sin \theta_n)$, $y_{n+1} = \mu(x_n \sin \theta_n + y_n \cos \theta_n)$, $\theta_n = 0.4 - \frac{6}{1+x_n^2+y_n^2}$. Color map (dark-light spectrum) according to the relative out-strength distribution of the directed multigraph W (which is equal to the Markovian stationary distribution of T). Ordinal patterns computed from the x-component time series with $m = 10$ and $w = 1$.

1. Discrete-time dynamics

Numerical results show that periodicity of a trajectory is captured by the ordinal symbolic itinerary s_n, although not necessarily at the original period. We present evidence that the corresponding OPN also preserves dynamical topology. Random walks from an arbitrarily chosen initial node recover the original symbolic trajectory with full certainty, provided a necessary and sufficient condition on m. If unfulfilled, then symbol sequences of lower period arise. We also make a preliminary case for selecting maximal overlap (slide lag $w = 1$) as it exhibits minimal sensitivity to the value of m, in contrast to all other overlapping variants.

In the trivial case of an equilibrium state, a time series comprises of a single repeated value. This results in a single-symbol itinerary regardless of chosen parameter values. The corresponding OPN constitutes of one node and a self-loop linking the node to itself. To consider regimes of longer period, we initially examine maximally overlapping symbolization.

Trajectories of period 2^k for $k \in \mathbb{N}$ lead to directed cyclic graphs of size 2^l where $l \in \mathbb{N}$ such that $0 \leq l \leq k$. The case $l = 0$ corresponds to the uninteresting $m = 1$. If $m \geq 2$, hence $l \neq 0$, then a random walk—irrespective of starting node—produces a 2^l-periodic symbolic itinerary since only one possible path exists within the network. This path is the network space analog of the original dynamical orbit. This is also the case for 3- and 5-periodic regimes, which leads to the hypothesis that orbits of any integer period may be captured by OPNs.

For instance, Fig. 8 depicts a visualization of the OPN ($m = 17$) computed from trajectories of the logistic map of period 2,3,4,8,16, and 32. We observe the respective directed 2^k-cycle (and 3-cycle) graphs, as expected. In fact, we shall show that a sufficiently high m allows the equality $l = k$ (and similarly for odd powers), in other words periodicity is captured *exactly* by the network topology. A random walk on this particular type of graph retrieves the original symbolic
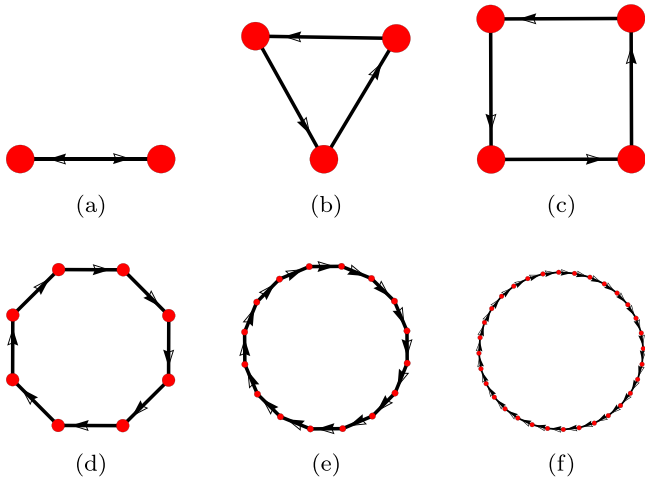
FIG. 8. OPN structure for $m = 17$ corresponding to trajectories drawn from (a) 2-periodic ($r = 3.4$), (b) 3-periodic ($r = 1 + \sqrt{8} \simeq 3.8284$), (c) 4-periodic ($r = 3.5$), (d) 8-periodic ($r = 3.56$), (e) 16-periodic ($r = 3.566$), and (g) 32-periodic ($r = 3.5695$) regimes of the logistic map. Maximal overlap between successive patterns ($w = 1$).

itinerary with probability equal to $\mathbb{P} = 1$. This finding is confirmed for time series generated by analogous parameter regimes (found within the period-doubling cascade regions in each) of the 1D cubic and the 2D Henon and Ikeda maps.

Figure 9(a) portrays the network size $V(m)$ for various values of m using maximal overlap. It is evident that the condition $m > 2^{k-1}$ is required to ensure the recovery of the original periodicity. Otherwise, the network's size will be equal to a submultiple of the original period and a random walk will produce symbol sequences of lower periodicity. This is due to the fact that an increasingly finer resolution of state space is necessary to encode the correct period. This signifies the requirement for an infinitely fine partition to fully encode the original dynamics if the accumulation point of transition to chaos (2^∞ periodicity) is under examination.
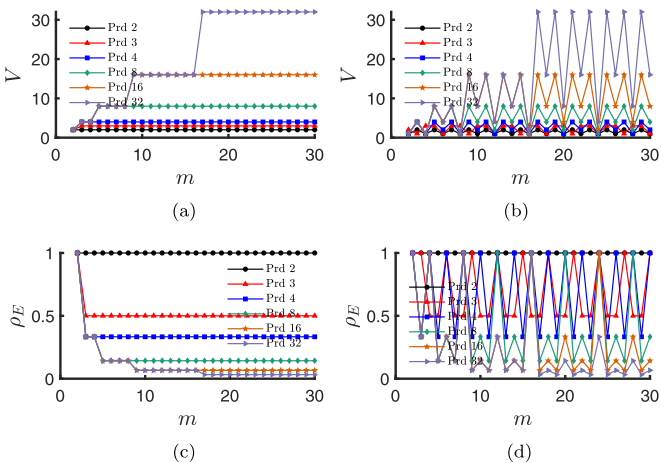

FIG. 9. Network size and link density as a function of pattern length m. Ordinal networks computed from the periodic trajectories of Fig. 8. Left: maximal overlap ($w = 1$). Right: zero overlap ($w = m$).

Such an unrealistic condition in practice motivates further the case for stochastic descriptions and Markovian modeling of low-dimensional dynamics. Additionally, there are two other comforting aspects. First, even if the detected periodicity is lower than the underlying, ordinal symbolization never fails to incorporate the periodic nature of the dynamics into the network topology. Second, for maximally overlapping symbols, OPNs exhibit no sensitivity to the value of m as long as it is selected above the threshold value of *half the original period*, i.e., $m = 2^{k-1}$ or larger. The existence of a limiting topology in the periodic case using maximal overlap is confirmed by looking at the corresponding link densities in Fig. 9(c). Dynamics are fully captured in this case.

The situation is radically different for the other extreme of the range of w, nonoverlapping OPN variants. Counting patterns with zero overlap leads to network topologies of varying size, as shown in Fig. 9(b), and link density, as portrayed in Fig. 9(d), which depend on an algebraic relationship between the underlying period and the length of ordinal patterns. Convergence to a limiting structure as m increases can, thus, never be established. There is no threshold value of m beyond which OPNs attain a constant structure. Both size and link density are fluctuating in an oscillatory fashion with increasing m, a fact that is true for all values of the slide lag $w \neq 1$. In fact, the network collapses to a single node if m is a whole multiple of the underlying period, with link density attaining the maximum allowed value of 1 (set to unity by convention since the network comprises a single node linked to itself).

2. Continuous-time dynamics

Encapsulation of periodic dynamics within an ordinal representation manifests in a different manner in the case of continuous-time flows. We discuss an example drawn from the Rössler equations. In particular, OPNs computed from trajectories of a periodic and a 4-periodic (Fig. 10) regime for various m do not lead to the directed k-cyclic structures observed for periodic regimes of discrete-time flows. Instead, resultant networks exhibit a 'ring'-resembling topology when m is large.

The reason for this type of imperfect—but simultaneously strongly recurrent, almost regular—topology is twofold. First, the continuous-time flows underlying the networks were sampled at approximately, and *not* exactly, 20 points per cycle ($T_{\mathrm{mean}}^{\mathrm{periodic}} = 6.1746$, $T_{\mathrm{mean}}^{4-\mathrm{periodic}} = 6.2078$ and $\Delta t = \frac{1}{20}T_{\mathrm{mean}}$). Therefore, integer m is required to be either a multiple of the mean cycle period T_{mean} or an exact factor (since $T_{\mathrm{mean}} \in \mathbb{R}$, by "exact factor" we mean a divisor such that the quotient is a positive integer). of T_{mean}. Second, comparatively larger values for m are necessary to reveal the original periodicity than in the case of discrete dynamics. For instance, were we able to sample the periodic regime at exactly 20 points per cycle, a value of $m = 20$ would be necessary to ensure that the corresponding network is (the correct) trivial structure of a single node and a single self-loop. One simple way to overcome this issue is by studying an appropriate Poincaré map, often referred to as Poincaré Surface of Section (PSS), of the continuous flow. We computed the x_{max} return map for the Rössler regimes using the *TISEAN* package [70] (here
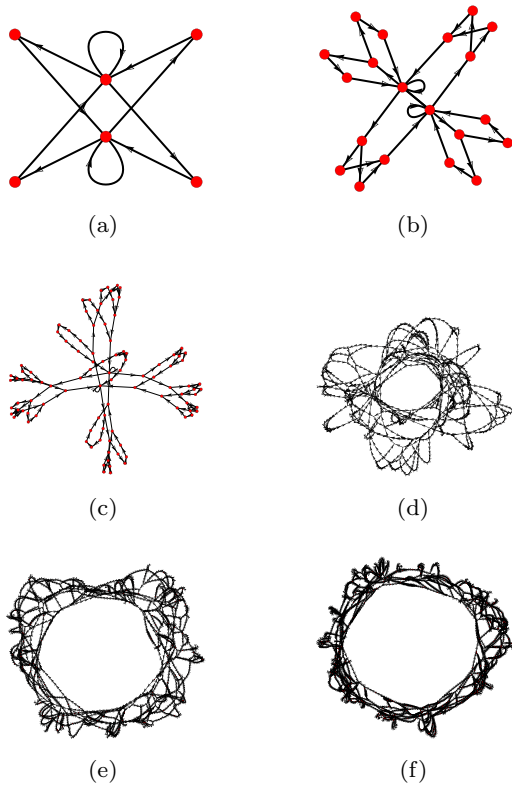
FIG. 10. Ordinal network structure using $w = 1$, $\tau = 1$ for a 4-periodic trajectory $(a, b, c) = (0.377, 2, 4)$ of the Rössler flow $\dot{x} = -(y + z)$, $\dot{y} = x + ay$, $\dot{z} = b + (x - c)z$ for (a) $m = 3$, (b) $m = 5$, (c) $m = 10$, (d) $m = 20$, (e) $m = 25$ and (f) $m = 30$. Mean cycle duration is $T_{\text{mean}} = 6.2078$ and sampling interval $\Delta t = \frac{1}{20}T_{\text{mean}}$.

the PSS corresponds to the x axis in tangent space [41]) to verify this. We observed that depending on precision of its values, an appropriate thresholding, i.e., a cut-off of decimal points exceeding the number of significant digits, leads to the recovery of the directed k-cyclic graphs observed for discrete-time flows.

B. Sensitivity analysis

Three methodological free parameters are involved in the construction of an OPN. The length of ordinal patterns m dictates the size of the symbolic alphabet (equal to $m!$), the level of resolution of the ordinal partition—which admits an inverse relationship with the alphabet's size (equal to $1/m!$)—and the dimension of embedding space. The time delay τ dictates the lag involved in the embedding transformation. Slide lag w is equal to the number of nonoverlapping points between successive patterns within the itinerary s_n. The main examples chosen for the purpose of our demonstrations are the logistic map [71] and the Lorenz flow [72] within well-studied parameter regimes (see Appendices B and C for the full collection of tested systems in discrete and continuous time).

The existing OPN studies suggest different procedures for selecting parameters. Small [7] propose using the graphical inflexion of the $V(m)$ curve or the maximum of the network link entropy (average node-link entropy) curve for identifying

the optimal value of m. The values $\tau = 1$ and $w = m$ are used for the other two parameters without providing strong reasoning. Sun *et al.* [26] performed their analysis with an arbitrary choice of $m = 8$ by arguing that networks computed with increasing m should in principle belong to the same class. The authors use $\tau = 1$ and $w = 1$ without providing a rationale behind this choice. McCullough *et al.* [19] suggest the interval $m \in [6, 10]$ as the most useful choice based on their empirical findings from testing model systems. The value for lag τ is selected based on the first zero of the autocorrelation function, whereas the slide lag $w = 1$ is picked without any motivation. Masoller *et al.* [27] select values based solely on the premise that large m will lead to undersampling. Hence, the arbitrary choice $m = 4$ is favored due to the small alphabet that it produces. The focus of this study is to examine short experimental time series, therefore no further consideration is given to determining parameter values and the choice $(\tau, w) = (1, 1)$ is not discussed. Similarly for Kulp *et al.* [10] who only propose a mechanism for selecting m by increasing it until a nonzero CFP is observed. McCullough *et al.* [9] vary τ, use $w = 1$ without discussing and pick m a posteriori by computing some metric on the network and observing its variation with increasing m. Zhang *et al.* [28] opt for $m = 2$ as its interpretation is simpler than higher values within the multivariate framework, $\tau = 1$ and $w = 1$ without supporting arguments. Guo *et al.* [29] do not discuss parameter selection.

In contrast, we make recommendations based on the sensitivity analysis that we conducted. Just as network visualization suggests (e.g., see structures computed with different m for the 4-periodic Rössler regime portrayed in Fig. 10), the value of m is of critical significance. A spectrum of values lying within a scaling region is appropriate. We, thus, recommend the ensuing analysis (Sec. V B 2) as a first step of application to a newly presented dataset. Furthermore, we posit that the set $(\tau, w) = (1, 1)$ is optimal based on arguments of (a) observing false admissible patterns and (b) exploiting short-term correlations to impose constraints on the admissible transitions. Additionally, we investigate the effects of undersampling and discuss facets related to this constraint.

1. Time delay

In this section we motivate the particular choice of the value $\tau = 1$ for the time delay, irrespective of the dataset at hand. In the case of discrete-time flows, the value of unity comprises a natural choice since sample points are essentially obtained at arbitrary precision (or, in practice, as far as numerical precision will allow). The first (and global) minimum of the mutual information between the time series and its history occurs at the value of 1 for discrete chaotic maps.

In the case of continuous-time flows, the decision is more complicated. If a longer time delay is required, then several heuristics for appropriate choice exist in the literature, most notably the first zero of the autocorrelation series (Ch. 1, Sec. 1.3 [73]) and the first minimum of the mutual information (Ch. 1, Sec. 1.4 [73]). System-dependent choices have also been proposed in the ordinal network literature [19]. Sometimes it is preferable to vary the time delay—especially when faced with multiscale dynamics—to identify the various time
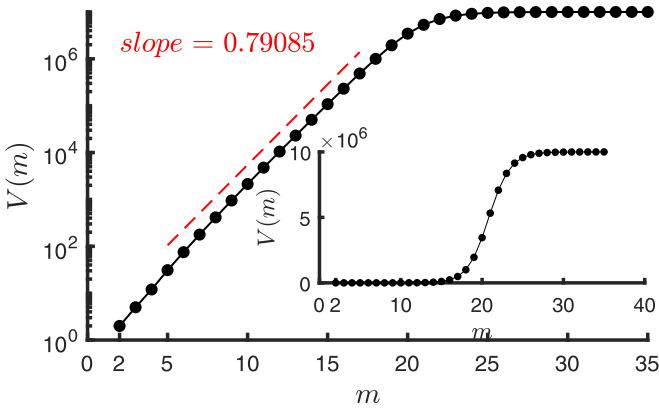
FIG. 11. Network size (total number of nodes) as m increases for $\tau = 1$ and $w = 1$ on semi-logarithmic scale axes. Ordinal patterns computed from the chaotic ($r = 4$) logistic trajectory. Inset: The same $V(m)$ curve on linear axes.

scales, as in the complexity-entropy causality plane analysis by Zunino *et al.* [74] or the ordinal analysis implemented by McCullough *et al.* [9] and applied to human cardiac dynamics.

We propose to opt for the $\tau = 1$ value via examination of an associated PSS. This renders the choice as natural as in the case of discrete-time flows. Our choice enables avoiding undesired aliasing effects [9], which is important to prevent the presence of *false* admissible patterns which do not reflect the actual deterministic dynamics. The absence of aliasing in itinerary s_n ensures the more accurate approximation of transition probabilities and connectivity in the resulting network.

2. Pattern length—Embedding dimension

Various ways for finding an optimal value for m have been suggested in the literature, most notably the *graphical inflexion* and *maximum network link entropy* propositions in the original paper [7]. McCullough *et al.* [19] suggested an alternative which considers the first two moments of the degree (vertex connectivity) distribution in combination to visual inspection. However, numerical evidence in this assay suggests that useful information may be obtained by considering various m values within an "optimal range." Our results indicate the existence of a *scaling region*. It is the equivalent of a combination of the scalings observed for various embedding dimensions and spatial scales ϵ in traditional embedding practice towards invariants' estimation. Therefore, we propose that instead of looking for an optimal value, it is more informative to focus on this region in parameter space and consider the interplay between mutually conflicting practical considerations, such as undersampling, on the one hand, and increasingly finer partitioning, on the other hand. All values of m within this region are satisfactory and then computational considerations may be the determining factor, i.e., picking the smallest possible value.

Figure 11 shows a plot of the total number of nodes for $2 \leqslant m \leqslant 35$ on a semilogarithmic scale. The corresponding OPNs were computed from a time series of $N = 10^7$ points generated by the logistic rule $x_{n+1} = 4x_n(1 - x_n)$. The inset figure on the bottom right portrays the $V(m)$ curve on linear axes. We observe a sigmoidal curve, whereupon it seems that

V increases slowly initially (for small values $m \leqslant 12$), then much faster until the inflexion point (intermediate values, $13 \leqslant m \leqslant 19$) is reached, subsequently the *rate of formation* of new nodes decreases ($20 \leqslant m \leqslant 24$) and, finally, a saturation occurs once m is sufficiently large ($m \geqslant 25$ here). The size of the network saturates to approximately the total number of time series points, $N = 10^7$ in this case.

The semilogarithmic plot shows that V is actually increasing exponentially fast for all m until the presaturation region $20 \leqslant m < 25$. Entering this region of parameter space signifies the transition from exponential to linear growth. Increasing m by a single unit produces one to two million additional nodes. We hypothesize that an OPN can provide a meaningful simplification to the underlying dynamics until some stage past the inflection point in the $V(m)$ curve. In the case of the chaotic Lorenz flow (sampled at discrete times, not the PSS), the inflection point usually occurs when m is slightly larger than an average period of the system. Although the rate of information incorporation into the network starts decreasing at this stage, some information about dynamical evolution may still be encoded within. Once $m > 25$, the network size exhibits gradual saturation to a constant value equal to $N - m + 1 \simeq N$, the total number of segments extracted from the time series. This transition signifies the collapse of the network representation to a single string. In this formulation, all nodes have strictly two connections to those representing the preceding and succeeding dynamical state, with the exception of the initial and final state which have one only. It is a completely trivial type of network without much use. Thus, once m represents a sufficiently large history of a trajectory in state space, the network is no longer a useful representation of the dynamics. The culprit is data length insufficiency. There is a limited amount of transitional information in the sense that far more connections between dynamical states occur due to the increasingly finer partitions but at a much lower frequency (recall that m affects the resolution limit $(m!)^{-1}$ in this type of partition).

To explain these observations, consider the pruning rule of forbidden patterns associated with deterministic systems [56]. Given infinitely long data and as m increases, the number of admissible patterns grows exponentially while the number of forbidden patterns grows at a super-exponential rate. Therefore, although the $V(m)$ curve is monotonically increasing, the relative count of admissible patterns $\mathcal{P}_a^{(m)}$ is actually decreasing rapidly. Finiticity of data renders this effect more pronounced, which implies we are possibly sampling only a portion of the real network. Undersampling occurs if m is larger than a certain threshold given fixed data length N. The heuristic condition $N \gg m! + m - 1$ is required to ensure an adequately large data set in comparison to a specified partition resolution $(m!)^{-1}$. Hence, up to $m \leqslant 10$ we have may sampled the network in full ($10! = 3, 628, 800$ and $N = 10^7$), while only partially for larger m.

Furthermore, admissible patterns grow exponentially as m increases. Bandt *et al.* [23] proved that for piecewise monotone one-dimensional maps the rate approaches the topological entropy of the evolution operator $h_{\text{TOP}}(\phi)$. In the case of continuous-time flows, this relates to the entropy of the discrete map sampled at regular intervals acting as the numerical approximation of the continuous system. Therefore, the
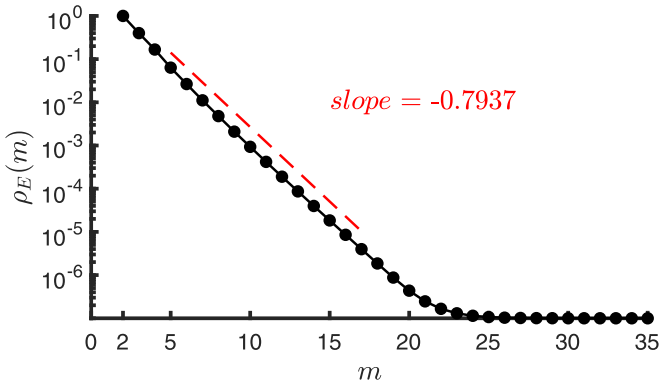
FIG. 12. Link density ρ_E as m increases ($\tau = 1$ and $w = 1$) on semi-logarithmic scale axes for the ordinal networks computed from the chaotic ($r = 4$) logistic trajectory.
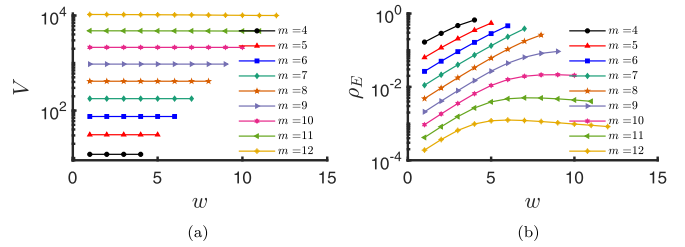

FIG. 13. Variation of network size (left) and link density (right) with respect to the slide lag between successive overlapping ordinal patterns. Networks computed for $4 \leqslant m \leqslant 12$ from the chaotic $r = 4$ logistic trajectory.

linear fit, $\frac{d}{dm}\{\log[V(m)]\}$, provides an estimate of $h_{\text{TOP}}(\phi)$, with a value approximately equal to 0.79085 for the logistic system. This can be contrasted to the actual value equal to $\log(2) \simeq 0.6931$. Disagreement between the linear fit and the true value also holds for Lorenz. In Ref. [34] we show that a network-based measure is better equipped to accurately capture this evasive dynamical invariant.

In summary, the domain of the parameter m contains three nontrivial qualitatively different regions in terms of the generated network topology—and a trivial one (saturation). First of all, if the pattern length m is not sufficiently large, then the size of the symbolic alphabet is too small to uncover the true complexity of the dynamics. This is akin to a histogram of very coarse resolution in traditional partitioning of time series data. The simplification is inadequate. Second, when m is larger than a critical value—presumably dependent on the dimensionality of the system and in particular the active degrees of freedom as established by Sauer *et al.* [61] in traditional embedding theory—the ordinal partition is finer and can capture the dynamical evolution of the system in a more elaborate manner. Very fine transitions from a dynamical state to another, which could not be captured by choosing smaller m, can now be incorporated into the network representation. In a practical setting, given finite data of possibly rather limited length (e.g., notoriously, geoscientific measurements), a heuristic rule consists of choosing as long a pattern length as the time series length will permit so that severe undersampling does not occur. Finally, within the third qualitative region, the network grows linearly with m and fewer newly formed nodes appear beyond the inflexion point. This point signifies that the difference between $V(m_{\text{inflexion}})$ and $V(m_{\text{inflexion}} - 1)$ is maximal. The ensuing decreasing rate of information incorporation into the network representation may then give rise to issues similar to those of over-embedding [7]. Due to this phenomenon and increasingly greater undersampling, the produced networks can no longer provide a meaningful simplification of the actual dynamics in play.

The above is further supported by looking at the link density of the computed networks which drops with increasing m, as depicted in Fig. 12. In fact, the semi-logarithmic scale plot indicates exponential decay. This result relates to the "existence" of *forbidden transitions* between symbols of

the alphabet associated with given deterministic dynamics. Such pruning concerns the product space of ordered pairs of patterns rather than the space of ordinal patterns itself. Note that the absolute value of the decay exponent in link density is very close to the growth rate exponent of the network's size.

An similar situation transpires for the nonoverlapping variant which suffers even more significantly for large m. A radical drop occurs at approximately $m = 25$. In both variants, for $m \geqslant 25$, we observe a decay to 10^{-7}, which occurs due to the summation term in Eq. (A1) approaching $V - 1$ since the network comprises a long string of $V = 10^7$ nodes connected by single directed links between predecessor-successor pairs of patterns. In fact, the large rise (and drop) in $V(m)$ ($\rho_E(m)$) discerned in Fig. 11 (Fig. 12) is also detected by measures such as PE, sorting entropy (SE; see Ref. [22]) and the entropy of the joint distribution of pairs of patterns in an analogous fashion (results omitted). The observations for $w = m$ networks, in contrast to $w = 1$, suggest a greater sensitivity displayed by the joint distribution of ordered pairs of patterns—in comparison to the more robust distribution of patterns—on the amount of overlap, as demonstrated in the following section.

3. Slide lag

Figure 13 depicts the two fundamental network observables for all possible overlap scenarios. It is evident that the effects on the network size V are negligible. However, measures which depend on the joint probability distribution of successive pairs of patterns exhibit sensitivity to the amount of overlap. This is highlighted by link density ρ_E in the right panel of Fig. 13. It increases with increasing slide lag, i.e., decreasing amount of overlap. This inverse relation implies an increasing diversity of forward-time transitions (and hence links in the network) as successive patterns become decreasingly overlapping and therefore less and less correlated.

Another interesting observation is the saturation of ρ_E beyond a certain slide lag. This may be explained by the fact that the number of forward-time transitions does not vary in total once the level of correlation between successive patterns has decayed to zero. This effect can only be observed for higher m values which allow complete decorrelation between consecutive symbol pairs. Since the value is similar irrespective of m once m is sufficiently large, it cannot merely be attributed to undersampling. This artefact has an interesting repercussion; it may be used as an indicator of the correlation time of a chaotic system and act as an ordinal alternative

to commonly employed estimates such as the first zero (or decay time) of the autocorrelation function—which is always zero after a single step for discrete-time flows examined here since there is no linear correlation, the first minimum of the self-mutual information and the quarter-period length.

Consequently, variation of w has no impact on the distribution of ordinal patterns, but it does affect the joint distribution of ordered pairs of successive patterns, on which correlation and/or undersampling effects are manifested in a pronounced manner with increasing m. While $w = m$ produces a richer diversity of admissible transitions, this is not necessarily beneficial. We therefore suggest the use of maximal overlap, thereby restricting the set of possible successor patterns. The networks produced are less sensitive to changes in m. Also, whereas the number of admissible patterns eventually observed is the same irrespective of w (see Sec. V B 4), reaching the true number of patterns for a *fixed* pattern length occurs for a much smaller sample size N in the case of the maximally overlapping variant. The conditioning on predecessor-successor pairs helps reduce undesired effects due to the Markovianity of the dynamics not being fully preserved in the projection to the symbol space—as with any dimensionality reduction procedure—unlike the embedding transform where topological conjugacy can be guaranteed.

4. Time-series length

It is self-evident that the longer a dataset, the more information an observer possesses about a system. In practice, however, measurements constitute small to intermediate-sized datasets in several disciplines. In the framework of ordinal patterns, small samples may lead to undersampling due to the super-exponential growth of the symbol space as m increases, e.g., 10! is approximately 3.6 million. Essentially, there is an interplay between finer resolution and insufficiently sampled regions in the ordinal partition $\mathcal{Q}$. The situation is even worse given that our aim is to estimate dynamical invariants from a single long trajectory rather than an ensemble of trajectories from randomly sampled initial conditions. This depends not only on the ergodicity of the underlying system, but also the capacity of the ordinal probability measure to preserve the ergodic property. The following results demonstrate that ordinal representations can prove faithful in the case of low-dimensional chaotic dynamics.

The relevant literature on the topic of undersampling mainly revolves around forbidden patterns and their outgrowth ratios [56] or PE, mostly in the context of white noise. An analytical treatise [75] illuminated the exact relation between PE of a Gaussian noise scalar time series to m and N, which obeys a χ^2-distribution of $(m - 1)!$ degrees of freedom. Several results on PE of fBM, fGN, and ARMA processes were obtained by Bandt and Shiha [76], but without any consideration of the dependence on the number of sample points N.

Here we are mainly interested in the effects of N when computing OPNs of increasing m from chaotic or hyper-chaotic dynamics. The most commonly employed heuristic for avoiding undersampling is to choose a value for m such that the sample size is significantly larger than the number of symbols, equal to $m!$. For maximal overlap ($w = 1$), this
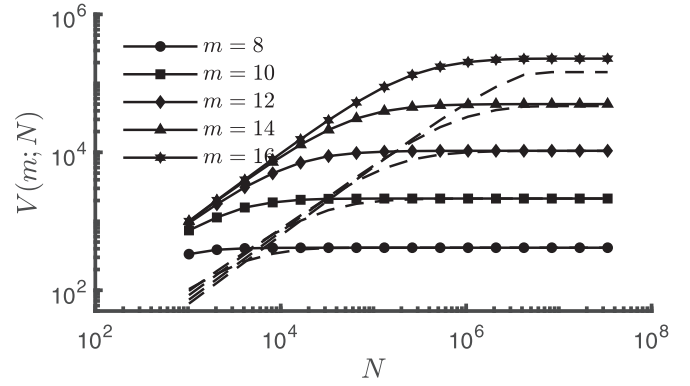

FIG. 14. Number of nodes vs. sample size N. Evolution of the network's size as the time series length increases. Network construction using maximal overlap ($w = 1$; solid lines with markers) and zero overlap ($w = m$; dashed lines) between successive patterns with $m = 8, 10, 12, 14, 16$.

leads to $N \gg m! + m - 1$ [56]. In the general case, $N \gg w(m! + 1) + m$ which has also led to the recommendation $N \gg 5m!$ [57,77]. However, our numerical experiments indicate that undersampling may be avoided even at a fraction of these sample sizes in the case of both dissipative as well as expanding low-dimensional chaotic flows.

Figure 14 shows the network size of OPNs computed from the ($r = 4$) logistic time series for various m as a function of N. When N is small, the network size increases linearly in the double-logarithmic scale of Fig. 14. This indicates a power-law relationship between sample size N and the number of distinct admissible ordinal patterns. In the maximal overlap case, the exponent is approximately equal to 1, i.e., there is a linear growth. In the nonoverlapping case, $V \sim N^{\frac{1}{10}}$, which shows slow growth and confirms the necessity for much longer trajectories (result omitted).

These findings are also evidenced by the saturation to limiting values (different for each m as expected), whereupon the maximal overlap variant attains a constant profile much faster. The surprising robustness to undersampling effects is apparent by considering $m = 8$ (circular markers) for instance. Although 8! = 40320, even approximately 2000 sample points are sufficient to obtain an accurate estimate of the network size. This is confirmed by looking at the corresponding link density in Fig. 14(b), whereby the saturation is now at lower values since the network becomes less dense *relative* to its size. For $m = 8$, a limiting link density is attained at about 3000 samples points. Finally, we remark that the differences resulting from the amount of overlap are perspicuous in these two fundamental structural network characteristics. There is, however, convergence between the two extremes of maximal and no overlap for adequately long time series.

C. Network topology

A comparison between networks computed from trajectories of different dynamics shows that a different topology is obtained—with the exception of when two regimes are topologically equivalent. A different size and link density is observed for the logistic map and time series from a chaotic cubic map, a Gauss map, the bent Baker's map,
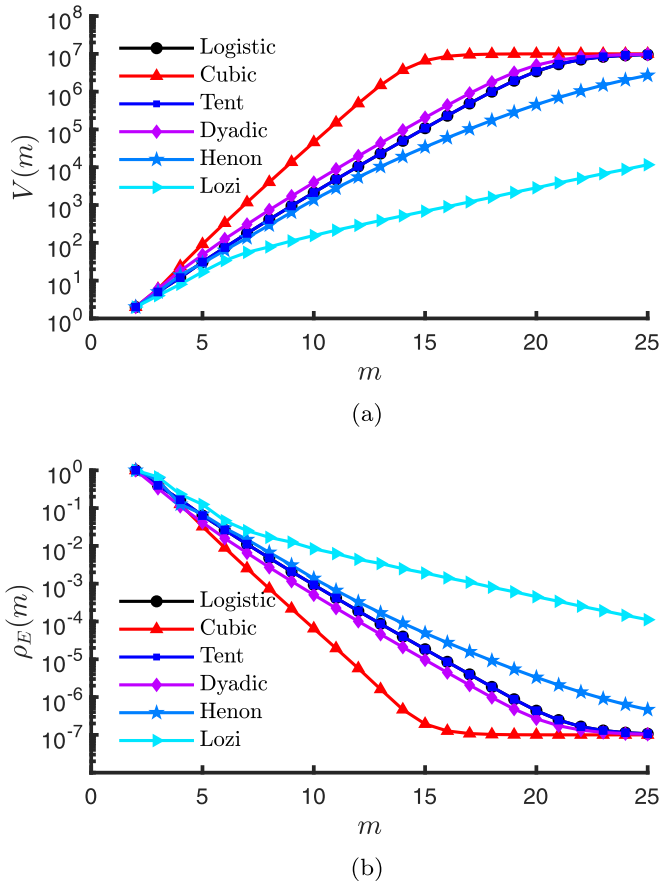
FIG. 16. Ordinal network computed with $m = 7$ from the first component of an $N = 1$-, 133-, 155-long 2D time series generated by the $z = 27$ PSS of the Lorenz attractor.

resemble the behavior that we observed for the logistic map. This is not so surprising in the case of the z_{max} PSS since it is very similar to the tent map (which we have also tested and obtained full agreement to these results) as the two discrete maps are topologically conjugate. However, the remarkable observation concerns the capacity of both types of PSS, the one-dimensional z_{max} as well as the two-dimensional $z = 27$, to generate very similar network topologies. Furthermore, any associated observable, for instance the corresponding x and y coordinates of the successive z maxima produce almost the same network. This result is further supported by exact quantitative estimates of the Lyapunov exponent [33] and the topological entropy [34] of the system. Unexpectedly, even series of recurrence times extracted from either type of PSS can also be used to procure similarly useful network topologies which are representative of the dynamics underlying the Lorenz attractor.

1. Out-strength distribution of directed multigraph W

In Sec. IV C we demonstrated an analytical connection between the node-wise relative out-strength distribution k_W^{out} of directed multigraph W and the stationary distribution of S. This connection is illustrated numerically for the logistic map example on the top panel of Fig. 18. First of all, observe the fat-tails characterizing this empirical distribution. A linear fit is shown on the inset logarithmic-scale plot. The slope is rather steep (estimated exponent $\gamma \simeq 4.21$ which is outside the typical range of $2 < \gamma < 3$), but the structure resembles

FIG. 17. (a) Network size and (b) link density as a function of m. Networks computed from seven observables generated by two Poincaré sections on the Lorenz attractor corresponding to the parameter regime $(\sigma, \rho, \beta) = (10, 28, \frac{8}{3})$. The seven time series comprise the (I) x and (II) y components of the $z = 27$ Poincaré map, the (III) x, (IV) y, and (V) z components of the z_{max} return map and (VI–VII) the time series of recurrence time for each map, respectively.

FIG. 15. (a) Number of nodes and (b) link density of networks computed from chaotic time series generated by various 1D and 2D discrete maps.

or the two-dimensional Lozi, Henon, and Ikeda maps. In contrast, the one-dimensional tent map and the dyadic (or bit-shift/doubling) map led to networks exhibiting similar structure. Computations on scalar x-component time series of a subset of these are displayed in Fig. 15. The exponential scaling behavior observed previously is apparent for all the maps considered, although at different rates. This, we conjecture, relates to the complexity characterizing the underlying dynamics. However, the width of the scaling region can be elongated in dynamics where the diversity of symbols is lower (smaller topological entropy), e.g., the chosen chaotic regime in the Lozi map exhibits a particularly long scaling region and markedly fewer patterns for respective m values. Its two-dimensional attractor is very thin with a particularly small Lebesgue measure.

We observed that similar conclusions hold for the two different types of PSS that we investigated in the case of the Lorenz attractor. Irrespective of the observable, all three components—and even the time series of the corresponding return times—produced similar network topologies. Figure 16 displays a visualization of the OPN computed with $m = 7$ from a $N = 1$-, 133-, 155-long trajectory of the $z = 27$ PSS of the Lorenz attractor (see Appendix C).

Evidently, spatial characteristics of the original attractor are present in the networks' structure depiction. The size of the network and link density as a function of m (Fig. 17)
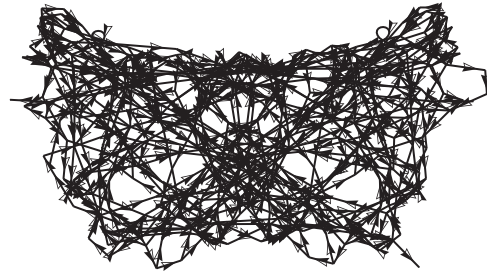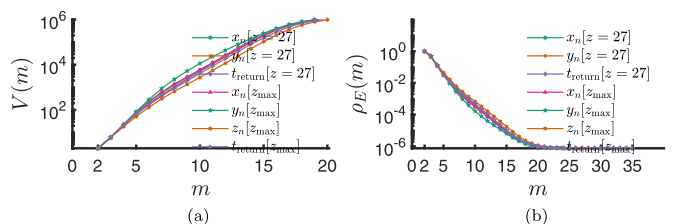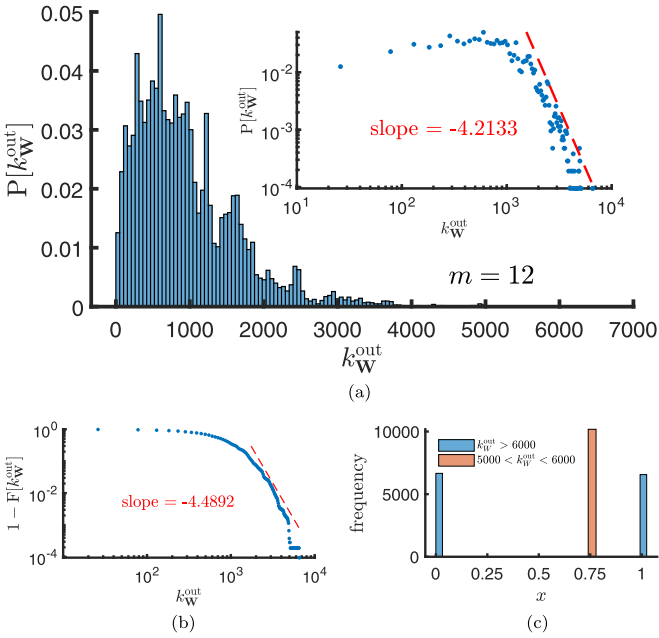
FIG. 18. (a) Relative out-strength distribution of directed multigraph W computed from the $N = 10^7$-long chaotic $r = 4$ logistic trajectory for $(m, \tau, w) = (12, 1, 1)$. Inset plot displays the corresponding logarithmic-scale plot. (b) Cumulative out-strength distribution function. (c) Number of time series segments $\boldsymbol{x}_i^{(m)}$ which have been mapped to the largest hubs, i.e., the nodes with the highest number of outgoing connections (k_W^{out} larger than 5 000 and 6 000); each class constitutes of only one node out of the approximately 10 000 total.

an empirical power law, i.e., a scale-free distribution. This observation is further supported by the associated plot of the cumulative distribution function, shown in the left-bottom panel.

We do not posit that OPNs computed from chaotic time series possess scale-free properties, as statistical analysis on empirical distributions has refuted many of the claims on reported scale-free networks and seriously questioned others [78]. In addition, simplistic mechanisms such as preferential attachment have absolutely no meaning within our context. Instead, we focus on the connection to the Markovian stationary distribution. It is likely that the scaling near the tails of the k_W^{out} distribution is a consequence of the singularities in the natural measure of the logistic map near the end-points of the domain. Since a higher concentration of mass occurs in these two regions of the one-dimensional state space, trajectories pass through this portion of space more frequently. Therefore, the symbolic itinerary s_n produced by an accurate embedding onto a higher-dimensional space should, in principle, reflect this property. This, in turn, would produce a higher number of adjacent links to the corresponding nodes of the network, i.e., those nodes which represent the relevant ordinal patterns in s_n—and by extension the elements of the partition $\mathcal{Q}$ of $\mathcal{M}^{\mathrm{IMG}}$.

Furthermore, the bottom-right panel of Fig. 18 reinforces this hypothesis. It displays the number of time series segments $\boldsymbol{x}_i^{(12)}$, which have been mapped to the two largest hubs of the network. As hubs we chose to collect the nodes whose number of outgoing connections k_W^{out} is larger than 5 000 (there are

two of them) and 6 000 (one of these two). Mapping back to the original state space, via labeling of the first coordinate of the 12-dimensional space, is illuminating. All trajectory points which the OPN algorithm has mapped to the partition element represented by the node with $k_W^{\mathrm{out}} > 6\,000$ are found very close to the end-points of interval [0,1], i.e., near the singularities of the invariant measure. Points mapped to the $5\,000 < k_W^{\mathrm{out}} < 6\,000$ node are found in the vicinity of the most dominant unstable periodic orbit (of order 1), i.e., the least unstable, which occurs at the fixed point $x = 0.75$.

2. Projection of ordinal partition onto 1D state space

To examine the above hypothesis in a more elaborate manner and understand the mechanism behind the proposed connection between k_W^{out} and the invariant density of the underlying dynamics, we present some indicative results on mapping back nodes to the original state space. Specifically, we illustrate the projection of a higher-dimensional ordinal partition back onto the original one-dimensional state space [0,1] of the logistic map for the intuitively simple and accessible cases of pattern length $m = 2$, 3, and 4.

All nodes of the resulting network have been projected back onto the original state space by inspection of the first coordinate of the two-, three-, and four-dimensional embedding space, respectively. The location of these points is shown on the horizontal axis of the three panels of Fig. 19, while the vertical axis represents the index of the node/pattern that these points correspond to, normalized by network size V so as to constrain the range within the interval [0,1]. These values are arbitrary and have been so chosen purely for visual purposes and facilitation of distinction between the different elements of the projection. For instance, the top-left panel portrays the projection of the $m = 2$ partition. There are two nodes, say v_1, v_2, corresponding to ordinal patterns π_1, π_2 and partition elements Q_1, Q_2 (shown in Fig. 4) of two-dimensional embedding space. In this case, $V = 2$ and index $i = 1$ or 2. Consequently, the vertical axis which represents i/V can take values from the discrete set $\{\frac{1}{2}, 1\}$. We have juxtaposed the zeroth, first, second and third iterates of the logistic function as well, respectively, on the plot of each panel.

Concentrating, first, on the $m = 2$ case, we notice that the "map-back" projection has separated the state space into two uneven intervals of Lebesgue measure equal to 0.75 for pattern π_1 and 0.25 for pattern π_2. The separation point occurs at the location $x = 0.75$, i.e., where the unstable fixed point lies. This is the point of intersection between the identity line $x = f^0(x)$ and the logistic function $f(x) = f^1(x)$. This is expected as a manifestation of the exchange of order relations between points in a single segment of length $m = 2$. On the left interval [0,0.75], we always have $x < f(x)$, and vice versa for all points contained in the interval [0.75,1].

The same observations can be extracted by inspection of the top-right panel which displays the projection of the $m = 3$ ordinal partition onto interval [0,1]. In this case, there are five intervals out of $3! = 6$ possible. The forbidden pattern $\{3, 2, 1\}$, which represents the order relation $f^3(x) > f^2(x) > f(x)$, corresponds to the yellow region of Fig. 5. Recall that this portion of embedding space can never be visited by the trajectory. Additionally, as in the $m = 2$ case, projected
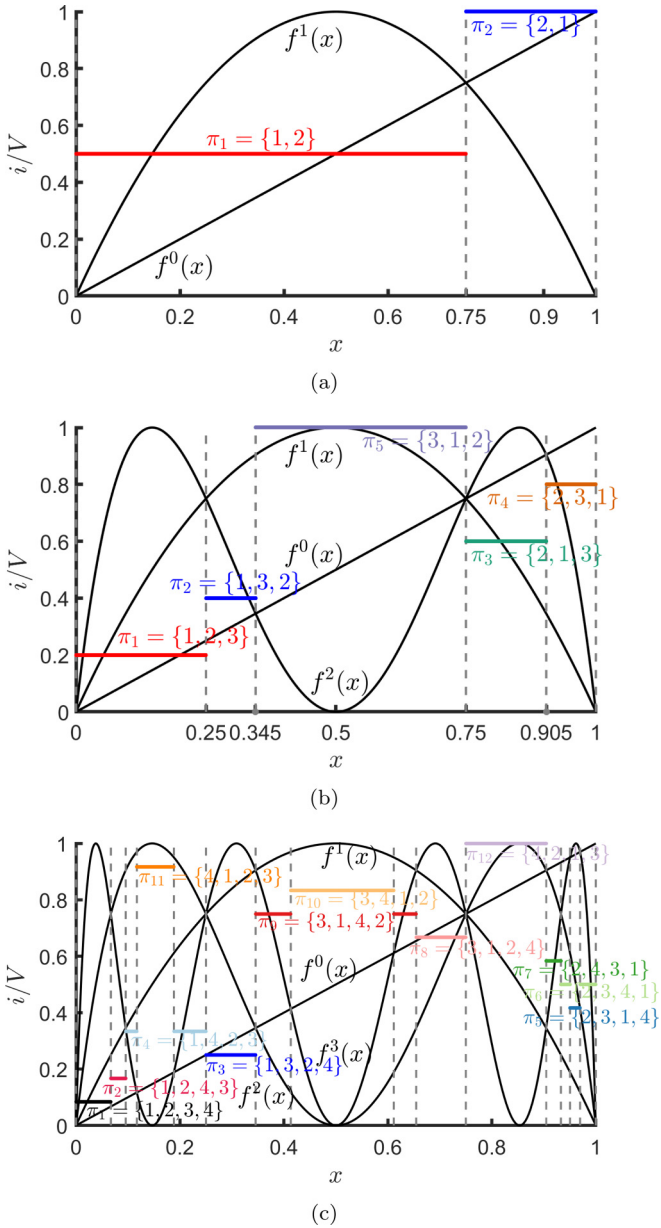
(a)

(b)

(c)

FIG. 19. Projection of the ordinal partition onto the one-dimensional state space of the logistic map. Top: $m = 2$ partition into 2 regions. Middle: $m = 3$ partition into 5 regions (recall that the pattern $(3,2,1)$ is forbidden). Bottom: $m = 4$ partition projection which consists of 12 elements. The vertical axis represents the index of each node $v_1, \ldots, v_{12}$ divided by the network's size $V = 12$ so that the range is bounded within $[0.1, 1]$. The first iterates $f^0(x), \ldots, f^{m-1}(x)$ are juxtaposed.

intervals have unequal length. Three of the four boundaries between these five intervals occur at the location of all the unstable periodic points of order 1 and 2. Namely, at the 2-periodic point $x = 0.3455$ whereby $f^2(x) = x$, at the 1-periodic (i.e., fixed) point $x = 0.75$ whereby $f^2(x) = f(x) = x$, and at the 2-periodic point $x = 0.9045$ whereby $f^2(x) = x$. The fourth boundary occurs at the eventually 1-periodic point $x = 0.25$ whereby $f^2(x) = f(x)$, which is mapped to 0.75 after one iteration. In this panel, the variable on the vertical axis can take values from the set $\{\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1\}$.
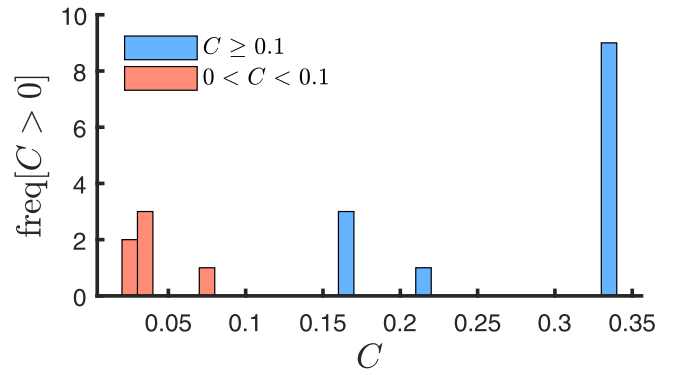

FIG. 20. Histogram of nodes with nonzero nodal clustering in the ordinal network computed with $m = 12$. Input: chaotic ($r = 4$) logistic trajectory.

Finally, the same trends are observed for the $m = 4$ partition projection, portrayed in the bottom panel of Fig. 19. Here there are 12 admissible patterns and another 12 forbidden ones. As previously, boundaries between elements of the projection coincide with either (a) periodic points of order 1, 2, and 3 or (b) eventually periodic points. An important observation is related to the length of each interval which, although unequal to the others, displays a decreasing length in comparison to length of the elements in the $m = 2, 3$ projections. This implies that the diameter of all elements of the ordinal partition projection onto the original state space decreases with increasing pattern length m. This is of significant value to the accuracy of the stochastic approximation, the Markov process S, to the underlying deterministic dynamics. Refinement of partition producing smaller elements in the Lebesgue measure constitutes a common requirement for traditional transfer operator approaches [30,31]. Note that one difference between the $m = 4$ case and the former two is that three elements of the projection, in particular those corresponding to the regions Q_4, Q_6, and Q_9 in four-dimensional embedding space—vertical coordinate equal to $\frac{4}{12} \simeq 0.333$, $\frac{6}{12} = 0.5$, and $\frac{9}{12} = 0.75$, respectively, in the bottom panel of Fig. 19—comprise the union of two disjoint intervals. This does not alter results or create any issues, but it does pose a peculiar property related to such partitioning schemes as with ordinal symbolization.

One final remark is that the inverse mapping procedure described above, can be used to identify the location of all periodic points of arbitrary order within a dataset. This also holds true for the location of all eventually periodic points. In contrast, if the underlying dynamics is dictated by a system of higher dimensionality, this procedure is no longer efficient for localising periodic orbits.

3. Nodal clustering

Very few nodes have nonzero local clustering coefficients (proportion of triangles over triplets passing through). For example, the OPN computed from the logistic map with $m = 12$ contains 10 525 nodes, while only 19 out of these do not exhibit zero clustering. Figure 20 displays the histogram of all such nodes, divided in two classes. The set of nodes with very low clustering ($C < 0.1$) appears in red, whereas
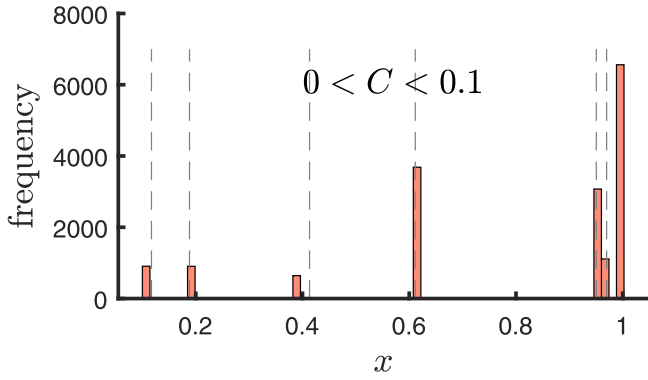
FIG. 21. Histogram of points that have been mapped to symbols corresponding to those nodes in the network with low clustering coefficient ($C < 0.1$). Gray dashed lines represent the 6 unstable periodic points of prime period 3.

FIG. 22. Number of k-cycles in the ordinal network as m increases.

nodes with higher clustering are shown in blue. We noticed that time-series points located in close proximity to the two fixed points of the map have been mapped to the latter class of symbols/nodes (blue).

In contrast, the former class (red) consists of symbols onto which points that lie close to the two 3-periodic orbits have been mapped. In fact, all six points with prime period 3 are encapsulated by this set of nodes. The histogram in Fig. 21 can be used to verify this as the gray dashed lines represent the exact location of all nondegenerate 3-periodic points.

This result is very indicative of the type of information encoded within the cyclic structure of OPNs. We delve deeper in the following section by considering cycles of arbitrary length. In summary, nodes characterized by nonzero clustering represent either fixed points of the underlying dynamics (higher values) or points with prime period 3 (lower values).

4. k-cycles in unweighted directed network A

Detecting the paths composed of k links from a node to any other within a network comprises a rather simple and efficient procedure. Given knowledge of the adjacency matrix, one need only look at the kth power of the matrix. Its (i, j) entry provides the number of all such walks from node i to node j. If specifically interested in the cyclic paths, i.e., those paths which end on the starting node whereby we have $j = i$, the so-called k-cycles, then the main diagonal of the kth power is required. Within the ordinal context, detecting k-cycles is rendered by the connectivity matrix, the adjacency matrix of the unweighted directed network A. A^k determines all the possible dynamical transitions from region Q_i to Q_j in embedding space within k time steps.

Figure 22 shows the base-2 logarithm of the number of k-cycles for the OPN computed from the logistic trajectory with $3 \leqslant m \leqslant 12$. It is clear that increasing m shifts the empirical curve closer to the diagonal, i.e., the expected curve since there exist 2^k unstable k-periodic points when $r = 4$. We utilize this result to detect the location of unstable periodic points in discrete maps via finer sieving obtained by increasing the pattern length.

We illustrate our results by considering the situation of $k = 4$ for the logistic map, which is amenable to an analytical
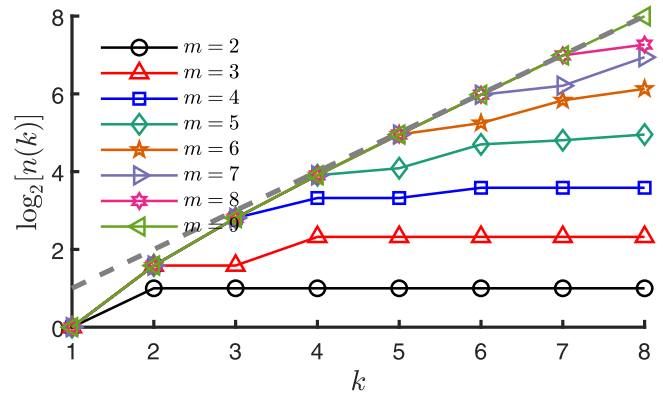
approach for all r values [71]. As a proof-of-concept for our conjecture, Fig. 23 displays a 64-bin histogram of those time series points which have been mapped to ordinal symbols belonging to 4-cycles in the computed network ($m = 6$). Gray dashed lines represent the 16 locations shown in Table II. It is evident that the vicinity of unstable periodic points of order $k = 4$ can be captured by the connectivity pattern of an OPN.

Increasingly finer sieving should, in principle, produce more precise results and shed light to the voluminous image portrayed in Fig. 23. As expected, Fig. 24 demonstrates that longer patterns ($m = 12$ here) are able to produce networks whose cycles encapsulate unstable periodic points embedded within a chaotic attractor—using the term slightly loosely here since the "attractor" in the logistic (and any expansive) map is composed of the entire state space, but see the following section for an example of a dissipative continuous-time flow.

The histogram of Fig. 24 illustrates three important facets of the described procedure. First, the aforementioned precision which implies that ordinal partitions of finer resolution allow one to detect unstable periodic points from a scalar time series with arbitrary precision. Second, the frequency (height) of each bin is dictated by the order of the periodic point, i.e., the lower the order, the more dominant the unstable periodic
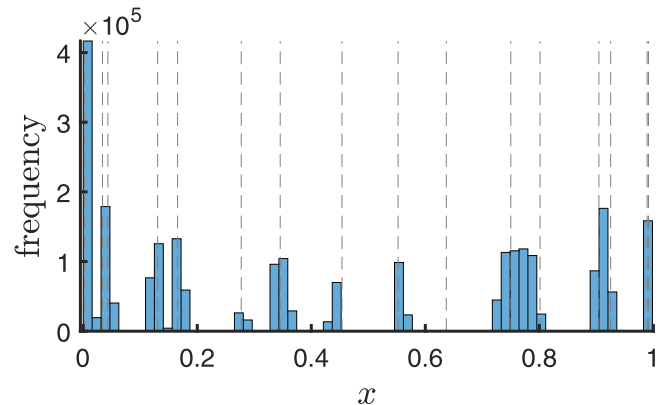

FIG. 23. Histogram of points (in 64 bins) that have been mapped to symbols corresponding to those nodes in the network ($m = 6$) which form 4-cycles. Computations drawn from the chaotic ($r = 4$) logistic trajectory.
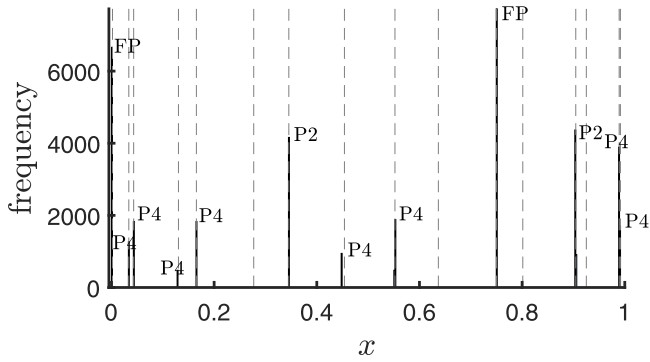
FIG. 24. Histogram (512 bins) of time series points mapped to network nodes which form part of 4-cycles for higher resolution ordinal partition ($m = 12$) than in Fig. 23. The location of all 4-periodic points of the logistic $r = 4$ chaos computed analytically are shown in gray dashed lines with indicative labels for fixed and periodic points.

point. Hence, more time series points may be found in the local neighborhood thereof. Finally, four points have been missed which indicates that a longer trajectory may be required to approach these periodic points sufficiently close and thereby allow the network to detect their location. Sec. V D illustrates the effectiveness of the proposed technique towards detection of *unstable periodic orbits* (UPOs) in continuous-time flows.

D. Detecting unstable periodic orbits in 3D chaotic flows

We use the Lorenz system as an example, in particular the two-dimensional $z = 27$ PSS. Ordinal networks were computed from the first component for various values of m and we focus on 4-cycles as previously. There are 18 such entries which are nonzero in the case of $m = 14$ which we use as a case study (Fig. 25). Collecting all the time series segments $\boldsymbol{x}_i^{(14)}$ mapped to the diagonal entries of $\boldsymbol{A}^4$, the plot displays a histogram of the first of the 14 coordinates in embedding space.
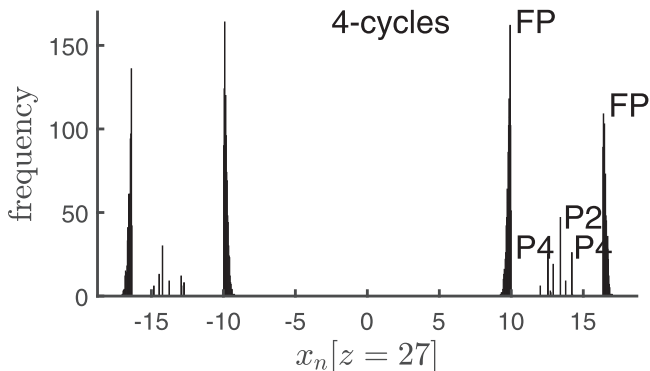

FIG. 25. Histogram of *first* component of segments $\boldsymbol{x}_i^{(14)}$ mapped to the 18 nodes (out of $V = 106\,586$) which form part of 4-cycles in the ordinal network computed from the x-component of the Lorenz $z = 27$ PSS. Peaks correspond to the two unstable fixed points and the lowest order unstable periodic orbits.
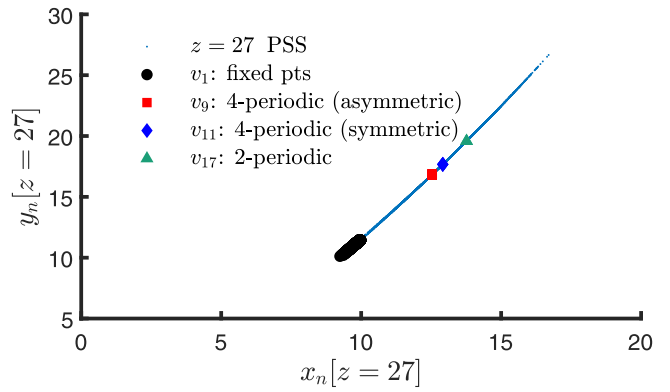

FIG. 26. Two-dimensional state-space plot of the positive part of the $z = 27$ PSS. Markers represent the location of time series points which have been mapped by the transform to the network nodes v_1, v_9, v_{11}, and v_{17} from the collection of 18 nodes which form part of 4-cycles. Points mapped to v_1 are located near the fixed point around which the right lobe of the Lorenz attractor is centered. Points mapped to v_9 and v_{11} lie close to a symmetric and an asymmetric 4-periodic orbit, while points mapped to v_{17} are located in the vicinity of the lowest order unstable 2-periodic orbit. See Fig. 27.

For comparison purposes (see Fig. 27 below), we have chosen to show frequency, and not its relative counterpart, on the vertical axis—as opposed to the results on the logistic map. Clearly, the symmetry of the Lorenz attractor is captured in this figure. We propose that the peaks represent the local neighborhood of the unstable fixed points around which the lobe dynamics of the attractor are revolving, as well as the lowest-order UPOs. Moreover, the highest peaks correspond to points of the PSS through which the elliptic trajectories near the fixed points transverse. The coordinates of the latter are, respectively, given by $x_{\mathrm{FP},1} = (\sqrt{\beta(\rho - 1)}, \sqrt{\beta(\rho - 1)}, \rho - 1) \simeq (8.4853, 8.4853, 27)$ and $x_{\mathrm{FP},2} = (-\sqrt{\beta(\rho - 1)}, -\sqrt{\beta(\rho - 1)}, \rho - 1) \simeq (-8.4853, -8.4853, 27)$. The second-highest peaks are due to points in the vicinity of the only UPO of order 2, the most dominant of all the periodic orbits. It is symmetric with respect to the z axis and somewhat resembles the ∞ symbol, shown in the inset figure of the bottom-right panel of Fig. 27. The next-highest peaks correspond to UPOs of order 4 and so forth. Evidence in support of our conjecture are provided via Figs. 26 and 27.

Figure 26 displays the right (positive) branch of the $z = 27$ PSS. Markers of four different colors and shapes have been placed at the location of all time series points that have been mapped to nodes v_1, v_9, v_{11}, and v_{17} from the set of the 18 nodes that form part of 4-cycles in the network. Nodes v_1, v_{18} are also identified as cycles of order 1, hence they correspond to the two fixed points. Nodes v_2, v_5, v_{15}, v_{17} form cycles of order 2. Therefore, we expect that v_1 encodes the neighborhood around fixed points, v_{17} the neighborhood around UPOs of order 2 and v_9, v_{11} the neighborhood around UPOs of order 4. Blue circles represent the points mapped to v_1, whose x and y coordinates may be compared to the first two components of $x_{\mathrm{FP},1}$. We conjecture that green triangular markers (mapped to v_{17}) correspond to points near the symmetric UPO of order 2, while red squares (v_9) and black diamonds (v_{11}),
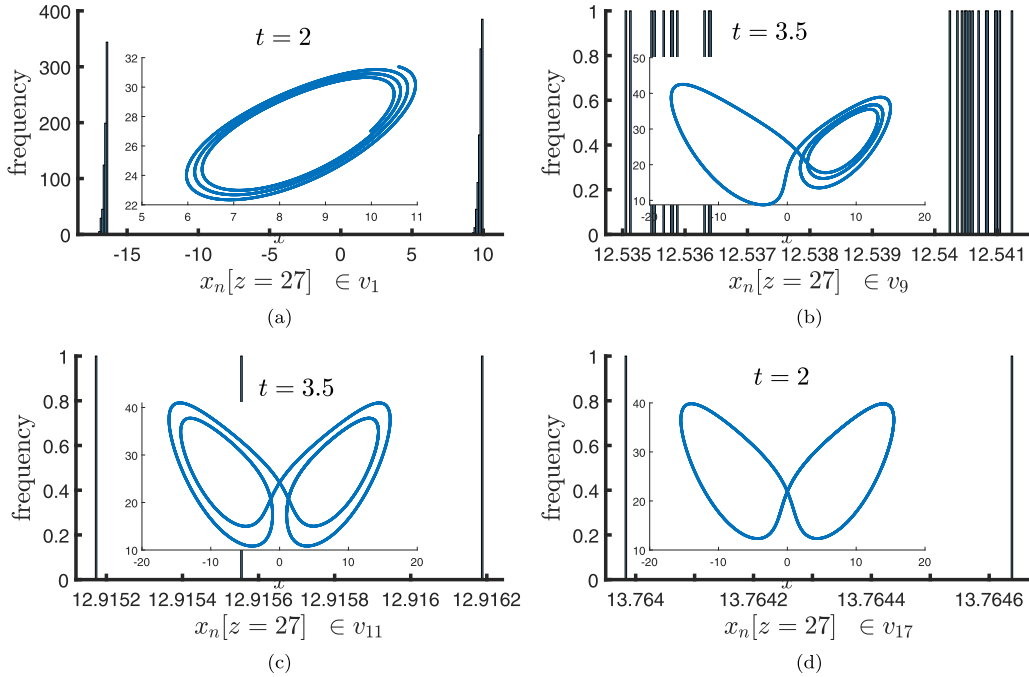
FIG. 27. Histogram of *first* component of segments $x_i^{(14)}$ mapped to nodes (a) v_1, (b) v_9, (c) v_{11}, and (d) v_{17} from the collection of 4-cycles in the ordinal network computed from the x component of the Lorenz $z = 27$ PSS. Inset plots: Numerical integration of example points in the histogram for a couple of cycles ($T_{\text{mean}} = 0.95757$).

respectively, represent the state-space points in the immediate neighborhood of one of the two asymmetric UPOs and the single symmetric UPO of order 4. To examine the validity of our hypothesis, we inspect individual histograms of the x component of all points mapped to each of the nodes v_1, v_9, v_{11}, and v_{17}. Furthermore, we have randomly sampled these 4 sets of points drawn from the $z = 27$ PSS, picked a single example point and numerically integrated the Lorenz equations for 2–3.5 time units.

The results of the procedure described above are presented in the four panels of Fig. 27. In the case of node v_1, there exists a multitude of time series points that have been mapped to this particular ordinal pattern, in the order of (10^3) or less (see the vertical axis of the histogram on the top-left panel of Fig. 27 which displays frequency). Since the sample size of the input time series is $N = 1, 133, 155$, these points form approximately 0.1% of the total. In stark contrast, points mapped to the other three nodes represent a very small proportion of the time series. Specifically, only 25 points have been mapped to v_9 (top-right panel), 3 to v_{11} (bottom-left panel), and only 2 to v_{17} (bottom-right panel). Additionally, notice how close the x coordinate (horizontal axis) of the points in each of these classes is to the rest of the members of the class. The inset figures in each panel demonstrate that the randomly sampled test points (one drawn from each class) are indeed extremely close to the aforementioned UPOS of order 4 (v_9 asymmetric and v_{11} symmetric) and the single one of order 2 (v_{17}). These results indicate that an OPN computed from a scalar time series is indeed capable of detecting the most-dominant UPOs embedded within the chaotic attractor of a continuous-time flow at sufficiently good precision.

VI. DISCUSSION

Ordinal partition networks comprise mappings from a scalar time series to a complex network. The input data is assumed to be generated by a deterministic dynamical source and our formulation is focused on ergodic systems. Network construction is conducted via a very simple, computationally efficient algorithm drawn from symbolic dynamics. It involves counting ordinal patterns, i.e., permutations of the numbers $1, 2, \ldots, m$ that represent order relations between points within short subsequences of the time series. The premise behind our proposition lies on a Markovian model which provides a stochastic approximation to the underlying dynamics. Our findings indicate that (i) the topology of the computed network encapsulates useful information about the dynamical features of the system and (ii) the mapping can be used to distinguish between different dynamical regimes.

In this work, we set out to test the extent to which information from the original time series is retained by the OPN transformation. Numerical experiments were conducted on an ensemble of systems and dynamical regimes. Our results suggest that the representation is faithful in the sense that connectivity patterns reflect characteristics of the dynamics and different dynamical regimes produce distinctions in the topological structure of the network. In addition, we only examined scalar time series and evidence suggests that any projection of a multidimensional flow can be used to extract a network representative of the underlying system. It can be deduced that any nonlinear monotonic transformation of any state variable will also exhibit the same ordinal structure. We presented an interpretation of OPNs that has a particular meaning within the context of ergodic theory and transfer operators. Through this lens, the network constitutes a natural

representation of a discrete-time finite-state Markov chain. This allowed us to relate the topology to either the periodicity or the hierarchy of unstable periodic orbits embedded within an attractor, in the case of periodic and chaotic dynamics, respectively. In addition, our framework sets the stage for using OPNs to estimate characteristic indicators of chaotic behavior [33,34].

Three different flavors of OPNs were introduced and all can provide insight into the governing dynamics. The relative out-strength distribution over all nodes of the directed multigraph W is equal to the stationary Markovian distribution, obtained by the ℓ_1 normalized left eigenvector of T. The transition graph T contains information about the transition probabilities between the elements of the ordinal partition in embedding space $\mathbb{R}^m$, where m the denotes the length of ordinal patterns. In Ref. [33], we illustrate the capabilities of this flavour in a quantitative manner by estimating the Lyapunov exponent of the chaotic system which has generated the time series.

The connectivity matrix A represents all possible transitions from any element in the ordinal partition of embedding space to any other. We utilized it to localize the most dominant (lowest-order k) UPOs of the underlying chaotic dynamical regime. This was achieved via identification of the time series points which have been mapped to cycles of order k in the network. This technique is effective both for discrete maps and continuous-time flow data by means of Poincaré sections. Any type of section that provides a sufficiently accurate depiction of the underlying dynamics may be used, even something as simple and computationally efficient as recording successive maxima from any component of a multidimensional system. Additionally, in Ref. [34] the authors show how it can be used to estimate the topological entropy of the underlying discrete-time flow. The above highlights the benefits of using this network-based technique over simply collecting statistics from the symbolic itinerary and the complementary insight obtained.

Finally, our aim being to provide a computationally cheap tool to augment traditional time-series analysis, we have created a guide for application of this toolkit to given data. Examination of methodological parameters suggests that the embedding lag (τ) and the amount of overlap between successive windows of the dataset segmentation ($m - w$) admit an optimal value. The former should be set to unity to avoid the aliasing effect, i.e., the occurrence of false admissible patterns that are not induced by the underlying dynamics. The latter parameter should be maximal, i.e., $w = 1$, to capture all possible information from the empirical joint distribution of time-ordered pairs of patterns. This leaves one free parameter, namely the pattern length (m) which also dictates the resolution of the ordinal partition, as well as the dimension of embedding space. Instead of an optimal value, we put forth the argument that there exists a scaling region, similarly to traditional embedding practice, where the resulting OPN captures the underlying dynamics sufficiently well. Within this region in parameter space, an OPN provides a useful stochastic approximation which can extract meaningful dynamical features from a scalar time series. An interplay between the sample size and the chosen value of m dictates the size of the optimal scaling region. Plotting the number of nodes as a function of

the sample size N—for given time series using increasingly larger subcomponents thereof—for various values of fixed m can be used to determine this region.

APPENDIX A: NETWORK METRICS

1. Network size

Denoted by $V = |\mathcal{V}|$, it is equal to total number of admissible patterns (permutations of order m) $\mathcal{N}(s_n, m)$ observed in the symbolic itinerary s_n.

2. Link density

Denoted by ρ_E, it constitutes the second fundamental network characteristic. Traditionally defined by

$$\rho_E = \frac{1}{V(V-1)} \sum_{i,j,\, i \neq j} A_{ij}, \tag{A1}$$

whereby self-loops (links from a node to itself) are not considered. This is equivalent to removing the main diagonal from adjacency matrix A. We have repeated the entirety of our analysis with a variant which includes self-loops. The normalizing constant changes to V^2 in this case. No differences in results were observed, with the exception of absolute density values.

3. Out-degree

Denoted by $k_A^{\text{out}}(i)$, it is defined as the total number of outgoing links adjacent to node i according to connectivity matrix A. Also referred to as *degree centrality*.

4. Out-strength

Denoted by $k_W^{\text{out}}(i)$, it comprises the extension of the out-degree metric to weighted networks. It equals the number of outgoing links from each node in the multigraph W, i.e., total number of transitions from a given pattern observed in itinerary s_n. In the case of the transition graph, we have $k_T^{\text{out}}(i) = 1$ by definition.

5. Clustering coefficient

The coefficient of nodal clustering is defined as the proportion of directed 3-cycles over all possible directed node triplets passing through a node in the network. In set builder notation it is given by

$$C(j) = \frac{|\{(i,k)\,:\,A_{ij} = 1 \cap A_{jk} = 1 \cap A_{ki} = 1\}|}{|\{(i,k)\,:\,A_{ij} = 1 \cap A_{jk} = 1\}|}. \tag{A2}$$

Note that the denominator equals $k_A^{\text{in}}(j) \cdot k_A^{\text{out}}(j)$, i.e., the number of serially connected 3-node subgraphs passing through j. The clustering coefficient of the network is the first moment of the distribution of C over all nodes, i.e., $C_{\text{clust}} = \frac{1}{V} \sum_j C$.

TABLE I. Dynamical behavior exhibited by iteration of the logistic map in different parameter regimes.

Regime	Stable node	Period 2	Period 3	Period 4	Period 8	Period 16	Period 32	Period 2^∞	Chaos
r	1.5	3.4	$1 + \sqrt{8} \simeq 3.8284$	3.5	3.56	3.566	3.5695	3.57	4

6. Closeness centrality

It represents the relative ease/speed of reaching any other node from a specified node in the network *on average*. It is defined as

$$CC(i) = \frac{1}{\sum_j d_{ij}}, \tag{A3}$$

where d_{ij} denotes the number of links that the shortest path from node i to j is composed of.

7. Eigenvector centrality

A natural extension of the concept of out-degree to include connections of length greater than unity. Within a social context and in undirected networks, it quantifies the indirect importance of a node via its links to other important nodes—a measure of influence. Within our setting, EC provides a measure of centrality for each symbol in s_n—hence, each element of partition $\mathcal{Q}$—by taking into consideration *arbitrary-length* transitions to other elements of $\mathcal{Q}$. This highlights basins of attraction according to the Markov model. It is given by the right eigenvector of the connectivity matrix A, i.e.,

$$EC(i) = \sum_j A_{ij} EC(j). \tag{A4}$$

8. *k*-cycles

Paths from a specified node to itself that consist of k links. This is equivalent to self-loops when $k = 1$ and bidirectional links when $k = 2$.

APPENDIX B: ITERATED MAPS

1. Logistic map (1D)

The one-dimensional (1D) logistic map with parameter $r \in [0, 4]$ is given by the equation

$$x_{n+1} = f(x_n; r) = rx_n(1 - x_n), \quad x_n \in [0, 1]. \tag{B1}$$

We focus on the $r = 4$ regime, whereby dynamics produce (i) expanding properties (true for all $r > 1$), i.e., volume-generating in the sense of the Lebesgue measure, (ii) chaotic behavior (true for all $r > r_c \simeq 3.57$), (iii) strong mixing properties, and (iv) the existence of 2^k unstable k-periodic points that are dense in the unit interval as $k \to \infty$. It constitutes a measure-preserving transformation of the 1D state space [0,1] to itself that is ergodic. The unique natural invariant measure, which is absolutely continuous with respect to the Lebesgue measure, is given by the density function

$$\mathbb{P}(x) = [\pi x(1 - x)]^{-\frac{1}{2}}, \tag{B2}$$

which admits two singularities near $x = 0, 1$.

In discrete-time flows, fixed points satisfy $x_{n+1} = x_n$ and, in this case, are located at $\text{Fix}(f; r) = \{0, 1 - \frac{1}{r}\}$. The origin exists $\forall r \in [0, 4]$ and is stable for $r \in [0, 1)$, while the nontrivial fixed point only exists (in a physically meaningful sense) for $r > 1$ and is stable for $r \in (1, 2)$. The set of periodic points occur when $x_{n+k} = x_n$. If they are of *prime* (or nondegenerate) period k, then we denote them by $\text{Per}_k(f; r) = \{x | f^k(x) = x, \; f^j(x) \neq x \text{ for } 0 < j < k\}$. The set $\text{Per}_2(f; r)$ consists of points which satisfy the equation $-r^2 x_n^2 + r(r + 1)x_n - (r + 1) = 0$, namely,

$$x_{2P,1} = \frac{1}{2} + \frac{1}{2r}[1 + \sqrt{(r+1)(r+3)}],$$

$$x_{2P,2} = \frac{1}{2} + \frac{1}{2r}[1 - \sqrt{(r+1)(r+3)}],$$

while $\text{Per}_4(f; r)$ is composed of points satisfying a 12th-order equation (omitted).

Table I shows a gamut of parameter values which generate inherently different dynamical regimes in the logistic map. Table II shows the exact locations of fixed points and periodic points of prime period 2 and 4 when $r = 4$.

2. Dyadic/bit-shift/doubling map (1D)

We examine the chaotic regime produced by iteration of

$$x_{n+1} = 2x_n \mod 1. \tag{B3}$$

3. Gauss map (1D)

We examine the chaotic regime for $(\alpha, \beta) = (6.2, -0.486)$ via iteration of the equation

$$x_{n+1} = e^{-\alpha x_n^2} + \beta. \tag{B4}$$

TABLE II. All 4-periodic points for the $r = 4$ logistic map.

Fix$(f; 4)$	Per$_2(f; 4)$	Per$_4(f; 4)$
0	0.3455	0.0338
0.75	0.9045	0.0432
–	–	0.1305
–	–	0.1654
–	–	0.2771
–	–	0.4539
–	–	0.5523
–	–	0.6368
–	–	0.8013
–	–	0.9251
–	–	0.9891
–	–	0.9915

4. Cubic map (1D)

We mainly examined the regime of "postmerge crisis" (PMC) chaos when $a = 4$, whereby the equation

$$x_{n+1} = (a-1)x_n - ax_n^3 \tag{B5}$$

generates behavior which admits the existence of an absolutely continuous measure (with respect to Lebesgue).

Regime	Period 2	Period 4	Period 8	Chaos	PMC Chaos
a	3.1	3.26	3.295	3.42	4

5. Tent map (1D)

We examined iterations of the equation

$$x_{n+1} = f_\mu(x_n) = \begin{cases} \mu x_n, & \text{if } x_n < \frac{1}{2} \\ \mu(1 - x_n), & \text{if } x_n \geqslant \frac{1}{2}, \end{cases} \tag{B6}$$

which produces chaos for $\mu = 2 - 10^{-8}$ and approximates behavior generated by f_2. Integer parameter values lead rapidly to convergence to the origin $x = 0$.

6. Lozi map (2D)

We examined chaotic behavior produced with $(a, b) = (\frac{6}{5}, -\frac{2}{15})$ via iteration of

$$\begin{aligned} x_{n+1} &= y_n, \\ y_{n+1} &= 1 + bx_n - a|y_n|. \end{aligned} \tag{B7}$$

7. Henon map (2D)

We examined the chaotic map given by

$$\begin{aligned} x_{n+1} &= 1 - ax_n^2 + y_n, \\ y_{n+1} &= bx_n, \end{aligned} \tag{B8}$$

in the traditional chaotic parameter set $(a, b) = (1.4, 0.3)$, as well as the Benedicks-Carleson chaotic regimes $(a, b) = (1.96, b)$ for $b > 0$ very small. The latter regimes have been shown to be ergodic and to generate a nonhyperbolic attractor which admits a unique SRB measure.

APPENDIX C: CONTINUOUS-TIME FLOWS

1. Lorenz system

The three-dimensional (3D) continuous-time flow generated by simulation of the Lorenz equations

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z, \end{aligned} \tag{C1}$$

is investigated using the original parameter values $(\sigma, \rho, \beta) = (10, 28, \frac{8}{3})$ which produce a "butterfly"-shaped fractal chaotic attractor. We compute two different types of a PSS. A 2D map can be obtained via recording the successive crossings of the Lorenz trajectory in the $\dot{z} > 0$ direction with the $z = 27$ plane. Three time series are stored, the x and y component of the location of the crossings, as well as the recurrence times. A one-dimensional Poincaré map can be obtained by the famous technique set forth by Lorenz [72] using the maxima of the trajectory in one direction, e.g., the sequence of successive maxima in the z component of the continuous flow. We store all three components associated with the $z_{\max}$ PSS, as well as corresponding return times.

2. Rössler system

We examined the system of equations given by

$$\begin{aligned} \dot{x} &= -(y + z), \\ \dot{y} &= x + ay, \\ \dot{z} &= b + (x - c)z, \end{aligned} \tag{C2}$$

mainly for $(a, b, c) = (0.398, 2, 4)$. This choice generates "broadband" chaos. Different regimes for $(b, c) = (2, 4)$ can be obtained by setting the value of a according to Table III.

3. Chua circuit

Another well-studied example is the three-dimensional system of equations used to describe the dynamics of a simple electronic circuit made from standard components only (resistors, capacitors, and inductors). It exhibits classic chaotic behavior for the set of parameters $(\alpha, \beta, m_0, m_1, E) = (9, 14 + \frac{2}{7}, -\frac{8}{7}, -\frac{5}{7}, 1)$. The resulting chaotic attractor is the so-called 'double-scroll' due to its shape. The system is given by

$$\begin{aligned} \dot{x} &= \alpha[y - x - \phi(x)], \\ \dot{y} &= x - y + z, \\ \dot{z} &= \beta y, \end{aligned} \tag{C3}$$

where the function ϕ—which describes the electrical response of the nonlinear resistor—is given by

$$\phi(x) = m_1 x + \tfrac{1}{2}(m_0 - m_1) \cdot (|x + E| - |x - E|). \tag{C4}$$

TABLE III. Dynamical regimes in the Rössler system.

Regime	Periodic	Period 2	Period 3	Period 4	Period 5	Period 8	Broadband	Fourband
a	0.3	0.35	0.41	0.377	0.4	0.3848	0.386	0.398

APPENDIX D: MATLAB CODE

```
1    function [adj] = OrdinalPartitionNetwork(xt,m,w)
2    %ORDINALPARTITIONNETWORK computes an ordinal network from a time series using
3    %patterns of specified length and overlapping segments of given lag
4    %
5    % FORMAT
6    % ======
7    % [adj] = OrdinalPartitionNetwork(xt,m,w)
8    %
9    %  INPUT
10   % =====
11   % xt: time series
12   % m: pattern length
13   % w: number of non-overlapping points between temporally successive segments
14   %
15   % OUTPUT
16   % ======
17   % adj: adjacency matrix of ordinal partition network
18
19   % length of input time series
20   N = length(xt);
21   % total number of segments
22   J = floor((N-m)/w) + 1;
23   % time series segmentation (embedding in m-dimensional Euclidean space
24   % using a single-point time delay)
25   Segments = zeros(m,J);
26   for j = 1 : J
27       Segments(:,j) = xt((j-1)*w+1:(j-1)*w+m);
28   end
29   % sequence of ordinal patterns corresponding to each time series segment
30   [~,s] = sort(Segments,1,'ascend');
31   % corresponding index sequence (database of distinct symbols)
32   [~,~,SymSeq] = unique(s','rows');
33   % sequence of single-step transitions (pairs of successive symbols)
34   Tij = [SymSeq(1:end-1) SymSeq(2:end)];
35   % distinct transitions
36   [Tij,~,idx] = unique(Tij,'rows');
37   % frequency of each disctinct transition
38   Tij(:,3) = histcounts(idx,'binMethod','integer');
39   % adjacency matrix (directed, weighted, with self-loops)
40   n = max(max(Tij(:,1)),max(Tij(:,2)));
41   adj = sparse(Tij(:,1),Tij(:,2),Tij(:,3),n,n);
42
43   end
```

[1] G. Nicolis, A. G. Cantu, and C. Nicolis, Int. J. Bifurcat. Chaos **15**, 3467 (2005).

[2] J. Zhang and M. Small, Phys. Rev. Lett. **96**, 238701 (2006).

[3] X. Xu, J. Zhang, and M. Small, Proc. Natl. Acad. Sci. USA **105**, 19601 (2008).

[4] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuno, Proc. Natl. Acad. Sci. USA **105**, 4972 (2008).

[5] Y. Yang and H. Yang, Physica A **387**, 1381 (2008).

[6] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths, New J. Phys. **12**, 033025 (2010).

[7] M. Small, in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)* (IEEE, Beijing, China, 2013), pp. 2509–2512.

[8] X. Gao, H. An, W. Fang, X. Huang, H. Li, W. Zhong, and Y. Ding, Phys. Rev. E **90**, 012818 (2014).

[9] M. McCullough, M. Small, H. H. C. Iu, and T. Stemler, Phil. Trans. R. Soc. A **375**, 20160292 (2017).

[10] C. W. Kulp, J. M. Chobot, H. R. Freitas, and G. D. Sprechini, Chaos **26**, 073114 (2016).

[11] J. Davidsen, P. Grassberger, and M. Paczuski, Phys. Rev. E **77**, 066104 (2008).

[12] N. Marwan, J. F. Donges, Y. Zou, R. V. Donner, and J. Kurths, Phys. Lett. A **373**, 4246 (2009).

[13] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, Europhys. Lett. **87**, 48007 (2009).

[14] Z. Gao and N. Jin, Phys. Rev. E **79**, 066303 (2009).

[15] J. B. Elsner, T. H. Jagger, and E. A. Fogarty, Geophys. Res. Lett. **36**, L16702 (2009).

[16] Y. Yang, J. Wang, H. Yang, and J. Mang, Physica A **388**, 4431 (2009).

[17] D. Eroglu, N. Marwan, S. Prasad, and J. Kurths, Nonlin. Processes Geophys. **21**, 1085 (2014).

[18] E. Zhuang, M. Small, and G. Feng, Physica A **410**, 483 (2014).

[19] M. McCullough, M. Small, T. Stemler, and H. H.-C. Iu, Chaos **25**, 053101 (2015).

[20] R. V. Donner, M. Small, J. F. Donges, N. Marwan, Y. Zou, R. Xiang, and J. Kurths, Int. J. Bifurcat. Chaos **21**, 1019 (2011).

[21] E. Bradley and H. Kantz, Chaos **25**, 097610 (2015).

[22] C. Bandt and B. Pompe, Phys. Rev. Lett. **88**, 174102 (2002).

[23] C. Bandt, G. Keller, and B. Pompe, Nonlinearity **15**, 1595 (2002).

[24] J. M. Amigó, M. B. Kennel, and L. Kocarev, Physica D **210**, 77 (2005).

[25] K. Keller and H. Lauffer, Int. J. Bifurcat. Chaos **13**, 2657 (2003).

[26] X. Sun, M. Small, Y. Zhao, and X. Xue, Chaos **24**, 024402 (2014).

[27] C. Masoller, Y. Hong, S. Ayad, F. Gustave, S. Barland, A. J. Pons, S. Gómez, and A. Arenas, New J. Phys. **17**, 023068 (2015).

[28] J. Zhang, J. Zhou, M. Tang, H. Guo, M. Small, and Y. Zou, Sci. Rep. **7**, 7795 (2017).

[29] H. Guo, J.-Y. Zhang, Y. Zou, and S.-G. Guan, Front. Phys. **13**, 130508 (2018).

[30] G. Froyland, in *Nonlinear Dynamics and Statistics*, edited by A. I. Mees (Birkhäuser, Boston, MA, 2001), Chap. 12, pp. 281–321.

[31] G. Froyland, Nonlinearity **12**, 79 (1999).

[32] M. Dellnitz and O. Junge, SIAM J. Numer. Anal. **36**, 491 (1999).

[33] K. Sakellariou, T. Stemler, and M. Small (unpublished).

[34] K. Sakellariou, T. Stemler, and M. Small (unpublished).

[35] Y. Zou, J. Heitzig, R. V. Donner, J. F. Donges, J. D. Farmer, R. Meucci, S. Euzzor, N. Marwan, and J. Kurths, Europhys. Lett. **98**, 48001 (2012).

[36] R. V. Donner, J. Heitzig, J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, Eur. Phys. J. B **84**, 653 (2011).

[37] Y. Hirata, M. Komuro, S. Horai, and K. Aihara, Int. J. Bifurcat. Chaos **25**, 1550168 (2015).

[38] A. Khor and M. Small, Chaos **26**, 043101 (2016).

[39] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, Phys. Rev. E **80**, 046103 (2009).

[40] K. Padberg, B. Thiere, R. Preis, and M. Dellnitz, Commun. Nonlinear Sci. Numer. Simul. **14**, 4176 (2009).

[41] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. (Cambridge University Press, Cambridge, 1997).

[42] J. P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).

[43] Y. Sinai, Dokl. Akad. Nauk SSR **124**, 768 (1959).

[44] G. Nicolis and C. Nicolis, Physica A **163**, 215 (1990).

[45] G. Froyland and K. Padberg, Physica D **238**, 1507 (2009).

[46] A. S. L. O. Campanharo, M. I. Sirer, R. D. Malmgren, F. M. Ramos, and L. A. N. Amaral, PLoS ONE **6**, e23378 (2011).

[47] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Applied Mathematical Sciences Vol. 42 (Springer, Berlin, 1983).

[48] K. Keller and M. Sinn, Nonlinearity **22**, 2417 (2009).

[49] Y. Cao, W. W. Tung, J. B. Gao, V. A. Protopopescu, and L. M. Hively, Phys. Rev. E **70**, 046217 (2004).

[50] G. Ouyang, C. Dang, D. A. Richards, and X. Li, Clin. Neurophysiol. **121**, 694 (2010).

[51] K. Schindler, H. Gast, L. Stieglitz, A. Stibal, M. Hauf, R. Wiest, L. Mariani, and C. Rummel, Epilepsia **52**, 1771 (2011).

[52] U. Parlitz, S. Berg, S. Luther, A. Schirdewan, J. Kurths, and N. Wessel, Comput. Biol. Med. **42**, 319 (2012).

[53] D. Cysarz, A. Porta, N. Montano, P. V. Leeuwen, J. Kurths, and N. Wessel, Eur. Phys. J. Special Topics **222**, 487 (2013).

[54] M. Barreiro, A. C. Marti, and C. Masoller, Chaos **21**, 013101 (2011).

[55] L. Zunino, M. Zanin, B. M. Tabak, D. G. Pérez, and O. A. Rosso, Physica A **388**, 2854 (2009).

[56] J. M. Amigó, S. Zambrano, and M. A. F. Sanjuán, Europhys. Lett. **79**, 50001 (2007).

[57] J. M. Amigó, S. Zambrano, and M. A. F. Sanjuán, Europhys. Lett. **83**, 60005 (2008).

[58] R. L. Devaney, *An Introduction to Chaotic Dynamical Systems* (Westview Press, Boulder, CO, 2003).

[59] G. D. Birkhoff, Proc. Natl. Acad. Sci. USA **17**, 656 (1931).

[60] F. Takens, in *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics Vol. 898 (Springer, Berlin, 1981), pp. 366–381.

[61] T. Sauer, J. A. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).

[62] J. M. Amigó, Int. J. Bifurcat. Chaos **19**, 3311 (2009).

[63] C. E. Shannon, Bell System Techn. J. **27**, 623 (1948).

[64] J. R. Norris, *Markov Chains*, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, Cambridge, 2009).

[65] D. Levin, Y. Peres, and E. Wilmer, *Markov Chains and Mixing Times*, 1st ed. (American Mathematical Society, Providence, RI, 2009).

[66] A. Khinchin, Uspekhi Mat. Nauk **11**, 17 (1956).

[67] T.-Y. Li, J. Approx. Theory **17**, 177 (1976).

[68] A. Boyarsky, J. Stat. Phys. **50**, 213 (1988).

[69] G. Froyland, Nonlinearity **12**, 1029 (1999).

[70] R. Hegger, H. Kantz, and T. Schreiber, Chaos **9**, 413 (1999).

[71] R. M. May, Nature **261**, 459 (1976).

[72] E. N. Lorenz, J. Atmos. Sci. **20**, 130 (1963).

[73] M. Small, *Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance*, Nonlinear Science Vol. 52 (World Scientific, Singapore, 2005).

[74] L. Zunino, M. C. Soriano, and O. A. Rosso, Phys. Rev. E **86**, 046210 (2012).

[75] D. J. Little and D. M. Kane, Phys. Rev. E **94**, 022118 (2016).

[76] C. Bandt and F. Shiha, J. Time Ser. Anal. **28**, 646 (2007).

[77] M. Riedl, A. Müller, and N. Wessel, Eur. Phys. J. Special Topics **222**, 249 (2013).

[78] A. Clauset, C. Shalizi, and M. Newman, SIAM Rev. **51**, 661 (2009).